Balancing Robustness and Redundancy in the Design of Environmental Sensor Networks

Setia Budi

School of Engineering and ICT
University of Tasmania

This dissertation is submitted for the degree of
Doctor of Philosophy

June 2018

Abstract

This thesis proposes a new approach in the design of Environmental Sensor Networks (ESN) by achieving the highest possible robustness with minimal redundancy. The proposed methodology produces the optimal number of sensor nodes require to best achieve its purpose, determines the optimal placement of sensor nodes, and investigates the impact caused by noise or gaps in the data. Noise and sensor data gaps are usually resulting from sensor errors (e.g., biofouling, electronics noise) or communication failures.

The distribution of sensor nodes in a given region is proposed using Evolutionary Algorithm (EA) as the optimisation tool. The main advantage of EA is the fact that it can test a large number of possible solutions without bias from local optima. The algorithm compares the best possible configuration of sensor nodes in an ESN using fitness function as the difference between the result yielded from the network and the historical data as a validated environmental models. The results obtained were promising, however, the proposed methodology relies on historical data. To overcome this limitation a set of mobile platforms (e.g., drones, animal-carrying sensors, robots, boats of opportunity) is simulated as collecting data from the environment (i.e., from a large modelling output set). The results of the mobile platform readings are then spatial-temporally interpolated and the results used by the EA to propose a configuration of the first ESN.

Validation for the proposed methods in this thesis is achieved by formulating and running the methods in a form of simulation study. The effectiveness of each ESN design produced in representing the RoI is compared against SouthEsk data model (i.e., as a representation for the actual measured value in the RoI). The performance of the proposed methods is also compared with some other methods in ESN design, including with expert knowledge.

The main contributions of this work are the measure of ESN representativeness which enable to assess and to compare the performance of different ESN designs; the method to find optimum ESN design which best represents a region of interest with a balance between minimum redundancy and maximum robustness; and the use of mobile platforms for data sampling to capture environmental behaviour in a region of interest which is useful for ESN design in the absence of historical data.

I dedicate this thesis to my beloved wife Lia Septiana.

Acknowledgements

First and foremost I would like to express my gratitude to my PhD supervisors for their passion and dedication in supporting and guiding me during these past three and a half years: Professor Paulo de Souza as my primary supervisor has been a great and resourceful mentor in guiding me on my journey in exploring the world of environmental sensor networks; Associate Professor Paul Turner for sharing his insight in scientific area and for the motivation boost during my candidature, Dr. Greg Timms for the wonderful quality times that we spent on discussing many technical details in my simulation studies; and Dr. Vishv Malhotra for sharing his insight in computational study and the time that he invested on verifying every mathematical notation that I formulated in my work.

My gratitude is extended to Sense-T and Commonwealth Scientific and Industrial Research Organisation (CSIRO) for providing me with a scholarship and top-up scholarship respectively. These scholarships enable me to complete my PhD as a full-time student. CSIRO also allows me to have access to their High Performance Computing (HPC) facility which enables me to run my simulation smoother and faster.

I would like to acknowledge the anonymous reviewers of the journal manuscripts (including CSIRO internal reviewers) for the priceless peer-review process which significantly improve the quality and the merit of the published articles, and as a consequence the quality of this thesis. Quality publications also required extra pair of eyes to assess the styling in the writing. In this case, I also acknowledge great support from Raymond Williams (CSIRO), Benita Vincent (CSIRO), Peter Marendy (CSIRO), Stephen Quarrell (Tasmanian Institute of Agriculture), Tom Gillard (University of Sydney), Peter and Josiah Kim (Maranatha Christian University) for investing their time in proof-reading my manuscripts and this thesis prior to submission.

My fellow PhD students for the friendship and valuable discussions throughout my candidature: Zhao Chen for his priceless wisdom and encouragement in my difficult times during my candidature; Ferry Susanto for each great time that we spent together in brain storming, planning and conducting simulations, writing journal articles, and heaps of fun stuff; Pascal Hirsch for being a wonderful office mate and for the great times that we spent together working on several micro Linux systems and RFID readers; Leandro Disiuta and

Oliverio Delgado for bringing colours in my PhD life with stories, jokes, and comfort. I wish you all the best for your PhD candidature and for your future path in science. May the Force be with you...

All the team members in Swarm Sensing project: Paulo de Souza, Benita Vincent, Stephen Quarrell, Geoff Allen, Peter Marendy, Auro Almeida, Dale Worledge, Andojo Ong, Ulrich Engelke, Huyen Nguyen, Raymond Williams, and Selim Mahbub for the great opportunity to be involved in a world class scientific project. A priceless opportunity which enables myself to collaborate with experts in many different areas.

My friends and family for providing support and comfort throughout my candidature. Finally my beloved wife Lia Septiana for her love and support.

Declaration

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Setia Budi
June 2018

Authority of Access

Setia Budi
June 2018

Statement of Co-authorship

The following people and institutions contributed to the publication of work undertaken as part of this thesis:

Name	Institution	Role
Setia Budi	School of Engineering and ICT, University of Tasmania	Candidate, Author 1
Paulo de Souza	Commonwealth Scientific and Industrial Organisation	Author 2
Paul Turner	University of Tasmania	Author 3
Greg Timms	Commonwealth Scientific and Industrial Organisation	Author 4
Vishv Malhotra	University of Tasmania	Author 5
Ferry Susanto	Victoria University	Author 6

The following are published works undertaken as part of this thesis, author details and their contributions are included :

[Manuscript 1] **S. Budi, P. de Souza, G. Timms, V. Malhotra, and P. Turner. Optimisation in the Design of Environmental Sensor Networks with Robustness Consideration. Sensors, 15(12):29765–29781, nov 2015. doi: 10.3390/s151229765.**

S. Budi (75%) contributed with the experimental work, data analysis, and manuscript writing. P. de Souza (10%) contributed in the experiment design, data analysis, and manuscript writing. G. Timms (5%), V Malhotra (5%) and P. Turner (5%) contributed in data analysis and manuscript writing.

[Manuscript 2] **S. Budi, F. Susanto, P. de Souza, G. Timms, V. Malhotra, and P. Turner. In search for a robust design of environmental sensor networks. Environmental Technology, 39(6):683-693, apr 2017. ISSN 0959-3330. doi: 10.1080/09593330.2017.1310303.**

S. Budi (70%) contributed with the experimental work, data analysis, and manuscript writing. F. Susanto (10%) contributed in the experiment design, data analysis, and manuscript writing. P. de Souza (5%), G. Timms (5%), V Malhotra (5%) and P. Turner (5%) contributed in data analysis and manuscript writing.

[Manuscript 3] **S. Budi, P. de Souza, G. Timms, F. Susanto, V. Malhotra, and P. Turner. Mobile platform sampling for designing environmental sensor networks. Environmental Monitoring and Assessment, 190(3):130–144, feb 2018. ISSN 1573-2959. doi: 10.1007/s10661-018-6510-0.**

S. Budi (70%) contributed with the experimental work, data analysis, and manuscript writing. P. de Souza (10%) contributed in the experiment design, data analysis, and manuscript writing. G. Timms (5%), F. Susanto (5%), V Malhotra (5%) and P. Turner (5%) contributed in data analysis and manuscript writing.

We the undersigned agree with the above stated “proportion of work undertaken"" for each of the above published peer-reviewed manuscripts contributing to this thesis:

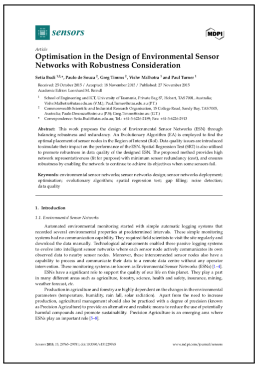
Setia Budi (Author 1)	Paulo de Souza (Author 2)	Paul Turner (Author 3)
Greg Timms (Author 4)	Vishv Malhotra (Author 5)	Ferry Susanto (Author 6)

Setia Budi
June 2018

List of Publications

The following are publications of work undertaken as part of this thesis:

1.

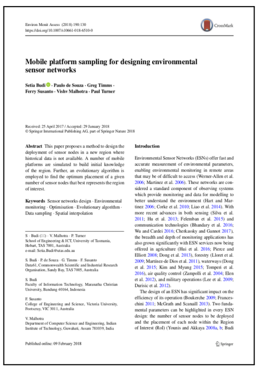


Title	:	**Optimisation in the Design of Environmental Sensor Networks with Robustness Consideration**
Authors	:	**Setia Budi**, Paulo de Souza, Greg Timms, Vishv Malhotra and Paul Turner
Journal	:	Sensors (MDPI)
JCR	:	2.677 (2016 Impact Factor)
Date	:	27 November 2015
DOI	:	**10.3390/s151229765**

2.

Title	:	**In Search for a Robust Design of Environmental Sensor Networks**
Authors	:	**Setia Budi**, Ferry Susanto, Paulo de Souza, Greg Timms, Vishv Malhotra and Paul Turner
Journal	:	Environmental Technology (Taylor & Francis)
JCR	:	1.751 (2016 Impact Factor)
Date	:	9 April 2017
DOI	:	**0.1080/09593330.2017.1310303**

3.



Title	:	**Mobile platform sampling for designing environmental sensor networks**
Authors	:	**Setia Budi**, Paulo de Souza, Greg Timms, Ferry Susanto, Vishv Malhotra and Paul Turner
Journal	:	Environmental Monitoring and Assessment (Springer)
JCR	:	1.687 (2016 Impact Factor)
Date	:	9 February 2018
DOI	:	**10.1007/s10661-018-6510-0**

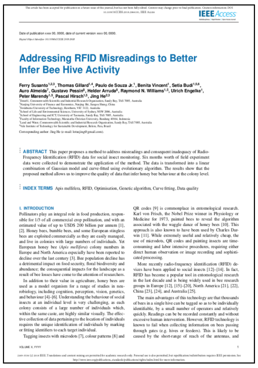
The following are publications of work undertaken within the period of PhD as co-author:

1.

Design of Environmental Sensor Networks Using Evolutionary Algorithms

Title	:	**Design of Environmental Sensor Networks using Evolutionary Algorithms**
Authors	:	Ferry Susanto, **Setia Budi**, Paulo de Souza, Ulrich Engelke and Jing He
Journal	:	Geoscience and Remote Sensing Letters (IEEE)
JCR	:	2.761 (2016 Impact Factor)
Date	:	26 February 2016
DOI	:	**10.1109/LGRS.2016.2525980**

2.



Addressing RFID Misreadings to Better Infer Bee Hive Activity

Title	:	**Addressing RFID Misreadings to Better Infer Bee Hive Activity**
Authors	:	Ferry Susanto1, Thomas Gillard1, Paulo de Souza, Benita Vincent, **Setia Budi**, Auro Almeida, Gustavo Pessin, Helder Arruda, Raymond N. Williams, Ulrich Engelke, Peter Marendy, Pascal Hirsch, Jing He
Journal	:	IEEE Access (IEEE)
JCR	:	3.244 (2016 Impact Factor)
Date	:	June 2018
DOI	:	**10.1109/ACCESS.2018.2844181**

3.

Low-Cost Electronic Tagging System for Bee Monitoring

Title	:	**Low-Cost Electronic Tagging System for Bee Monitoring**
Authors	:	Paulo de Souza, Peter Marendy, Karien Barbosa, **Setia Budi**, Pascal Hirsch, Nasiha Nikolic, Tom Gunthorpe, Gustavo Pessin, Andrew Davie
Journal	:	Sensors (MDPI)
JCR	:	2.677 (2016 Impact Factor)
Date	:	Minor Revision
DOI	:	**10.20944/preprints201805.0236.v1**

4.	Title	:	**Agent-based Modelling of Honey Bee Forager Flight Behaviour for Swarm Sensing Applications**
	Authors	:	Paulo de Souza, Raymond Williams, Stephen Quarrell, **Setia Budi**, Ferry Susanto, Benita Vincent, Geoff Allen, Auro Almeida, Dale Worledge, Leandro Disiuta, Pascal Hirsch, Gustavo Pessin, Helder Arruda, Peter Marendy, Leon dos Santos, Tom Gillard and Andojo Ongkodjojo Ong
	Status	:	**Under Review**
	Journal	:	Environmental Modelling & Software (Elsevier)
	JCR	:	4.404 (2016 Impact Factor)
5.	Title	:	**Data-driven Field Simulation and Environmental Modelling for Swarm Sensing Project**
	Authors	:	Ferry Susanto, Paulo de Souza, Raymond Williams, **Setia Budi** and Peter Marendy
	Status	:	**Under Review**
	Journal	:	Transactions on Geoscience and Remote Sensing (IEEE)
	JCR	:	4.942 (2016 Impact Factor)

Table of contents

List of figures

List of tables

Chapter 1

Introduction

Human activity, in the social, economic, health, and safety aspects is impacted by the states of the environment. Favourable environmental conditions contribute in delivering the best outcomes for human activity. Meanwhile, additional regulations are required to deal with harsh environmental conditions to minimise the adversity caused by unfavourable events in the environment (e.g., bush fires, flood, drought). A good understanding of the environmental phenomena (e.g., what happened, when did it occur, what was the magnitude and the duration of the event, how far did its effects spread) will benefit the regulation of human activity. The past and recent records of environmental parameters (e.g., air temperature, relative humidity, wind speed, rain fall, solar radiation) are essential to better understand the environment. In addition, environmental records from a well-monitored region could be used to model the region and forecast its environmental states.

Environmental Sensor Networks (ESN) extend the ability of human to measure and record environmental parameters by greatly increasing the frequency and representation of the measurements, as well as their accuracy [68, 121]. These networks are crucial in supporting informed decision-making in businesses and communities impacted by environmental changes.

Automated environmental monitoring began with simple automatic logging systems that periodically recorded a number of environmental properties. The lack of communication capability in the early monitoring systems required field scientists to visit the site regularly and collect the recorded data manually [32, 37, 110].

Over the past decades, there have been a number of technological advances in sensing instrumentation [23, 55, 74, 135], data transmission [13, 153, 96], data formatting and interoperability [125, 126], data management and information storage [102], as well as data processing and analytics [104, 105] which transformed the passive logging systems into intelligent sensor networks. These interconnected sensor nodes are capable of processing and communicating their data to a remote data centre without any operator intervention. The networks offer considerably faster and more accurate measurements, compared to the manually accessed logging systems, especially in remote areas which are difficult and sometimes risky or too expensive to reach. These capabilities enable scientists and decision makers to have high quality environmental sensing data which is fundamental for forecast modelling and better decision-making when environmental parameters are relevant.

Predicting changes in the environmental parameters over an extensive region is vital for a number of activities including agriculture and forestry [71, 118, 120], water and air quality assurance [47, 133, 137], logistics [10, 114, 134], tourism and recreation [63, 82], urban development [54, 98, 146], and emergency responses [50, 61]. The application of ESN in these areas will have a significant role to play in improving the quality of human life.

Agriculture and forestry are highly influenced by the changes in the environmental parameters. Agricultural production could be significantly improved by utilising agricultural management strategies with a greater degree of precision (i.e., Precision Agriculture). Precision Agriculture is an emerging area where ESN plays an important role [14, 56]. Apart from the need to increase production, the application of ESN in agriculture and forestry could provide an alternative and realistic means in minimising the use of potentially harmful compounds (e.g., insecticides) in the environment and promoting sustainable agricultural practices. ESN has also been successfully utilised for fire detection systems in forestry [108, 112, 144]. The networks can alert emergency services in the initial detection of the fire before it has spread uncontrollably and destroys hectares of vegetation incurring social, environmental and economic costs. The need to promote better healthcare also motivates the extensive use of ESN to monitor water quality [47, 86, 137] and air quality [51, 109, 133].

In science, ESN not only enables us to find answers to many scientific questions (related to our environment) which could not be answered in the past, but also prompts new questions which have yet to be asked. Furthermore, without ESN, the changes in the earth's climate would never be readily identified. ESN has been envisioned as a standard component in earth system and environmental sciences which enables scientists to better understand the environment and its phenomena [37, 68].

The design of an ESN has significant impact on the efficiency of its operation [17, 60, 113]. Two fundamental parameters can be highlighted in every ESN design: the number of sensor nodes to be deployed and the placement of each node within the Region of Interest (RoI) [158, 139, 20]. The complexity of this design problem increases with a requirement to have a fully functional ESN with the least possible number of sensor nodes [44, 107]. Within the current study, the decision-making process in relation to these two parameters are mainly shaped by the past measurements (i.e., historical data) conducted in the RoI. This historical data is utilised to capture the general characteristics of a particular environmental parameter over a certain period of time (e.g., seasonal, annual) [163, 95, 33].

The work presented in this thesis is aimed at contributing to the design aspects of ESN. This work seeks to find the best configuration for sensor nodes with the balance between redundancy and robustness in the networks. Historical records are valuable resource for ESN design, as ESN deployment are most likely to be an evolutionary development. A set of manually monitored, non-integrated sensors may be replaced by an integrated ESN with the designed configuration. In the situation where some of the sensors become non-operative (i.e., sensor failure), there is a necessity to maintain the data quality of the ESN. A graceful degradation in ESN performance is expected. The exploration for the work in this thesis also covers an effort to overcome the unavailability of past measurement records to support the design of ESN. An efficient method is needed to cover a geographical region and gather adequate data in the region that has not been previously monitored.

The following sections in this chapter will present the motivation which signifies the importance of the study, followed by research questions and research objectives that will guide the study presented in this thesis. The overall structure of the thesis will be presented in the last section of this chapter.

1.1 Motivation

Design is a critical process prior to the deployment of Environmental Sensor Networks (ESNs); it is one of the most significant factors in ensuring that the network delivers fit-for-purpose data and is cost-effective. There are two important questions in ESN design:

- How many sensor nodes are needed to serve they serve their purpose as an ESN?
- Where each sensor node should be deployed in the RoI?

The combination between the number of sensor nodes and the size of the RoI will significantly impact the number of possible placement of the nodes in the RoI, where each placement yields different level of representativeness. Complexity is introduced especially when dealing with the requirement to have a fully operational ESN, which meets the application purposes, with the lowest possible number of sensor nodes [19].

Apart from achieving a particular deployment goal, minimising the costs is the main focus in ESN design. In the current practice of ESN design, one of the major foci for reducing costs is to optimise the use and total number of sensors required to cover a specific area of interest [158, 159]. In most cases, redundancy in the ESN design is considered inefficient and should be reduced.

Deploying complex equipment like a sensor node in a remote environment is a very challenging task. There are numerous possible events, starting with harsh environmental conditions to animal activities, which may lead to sensor failure. As sensors start to fail, the usefulness of the network degrades. If an ESN no longer produces the data needed; it is not advisable, or even possible, to rely on data from such a network for decision-making [36, 123, 162]. Having an effective and fully operational ESN is costly and difficult to maintain. Minimising operational costs while delivering useful information is a constant trade-off. In principle, a robust ESN can be achieved by over sampling, at a potentially prohibitive cost. Nevertheless, redundancy in sensor node deployment would also introduce an undesirable increase in costs (e.g., deployment and maintenance costs), which is considered as inefficient in most design practices. For this reason, it is important to find a compromise between ensuring maximum robustness (i.e., fit-for-purpose) and minimising redundancy (i.e., cost-effective). Finding this balance is an optimisation problem.

ESN design is crucial and many studies have been carried out in this space which include several classical parameters such as the number of sensors required to adequately cover a specific area, the position of these sensor nodes, and the required frequency of readings and period of deployment. However, quality assurance and quality control (QA/QC) are neglected in most of ESN design practices. For a balanced design methodology, it is essential to identify an optimum number of sensor nodes (including the placement of each node) which best represents the RoI with low level of redundancy without sacrificing the robustness of the network. Balancing the robustness and the redundancy in the design of an ESN is an interesting yet challenging research problem. The work reported in this thesis is focused on optimising the design of ESN, with particular consideration in the quality of the data supplied by the network. This should result in an ESN that minimises its redundancy (i.e., cost effective) while maintaining the robustness of the network (i.e., greater trust in the data).

1.2 Research Questions

In order to give a clear direction on the study conducted in this thesis, four research questions have been formulated:

- **Q1: How to determine the minimum number of sensor nodes to be included in an ESN design?**
  In ESN design, the decision on the number of sensor nodes to be deployed has a direct influence on both the deployment and the operational costs of the network. An excessive deployment of sensor nodes may introduce an unnecessary increase in the deployment cost and also lead to inefficiency in its operation. A systematic method to

assist such decision making is essential in order to promote efficiency in ESN design. This research question will guide the study conducted in this thesis in exploring and reviewing the current practices in ESN design in determining the number of sensor nodes. The acquired knowledge can be used as a good foundation to construct and to propose a new method to address the question.

- **Q2: How to determine the placement of sensor nodes in an ESN design?**
  The effectiveness of an ESN design is highly influenced by the the placement of each sensor nodes within an RoI. ESN data is expected to represent the region where the network is deployed. Arbitrary placement of sensor nodes may result to ESN data which fails to represent the RoI. The measured data could be completely unusable and thus the network failed to serve its purpose. A systematic method to find an optimum placement of sensor nodes is needed. This research question paired with Q1 would serve as a guidance for this study to formulate a method in finding an optimum node placement. Such optimum placement would result to an ESN data which best represent the region and eventually promotes the effectiveness of the network.

- **Q3: How to improve robustness in an ESN design?**
  Efficiency is the major focus in current ESN design practices, which mainly deals with minimising the redundancy in the network. However, sensor node failure is common in ESN (i.e., many factors such as harsh environmental conditions could lead to sensor failure). This situation may lead to a condition where an ESN is no longer able to produce data which serves its purpose. Therefore, a certain degree of redundancy might be helpful to preserve the robustness of the network. Finding a balance between redundancy and robustness is needed. This third research question will guide this study to propose a method which incorporates robustness in the design of ESN.

- **Q4: How to design an ESN in the absence of historical data?**
  Designing an ESN for a new region, where no previous environmental monitoring has been conducted, is a challenging task. The absence of historical records as reference makes the decision on assigning number of sensor nodes in a new region difficult. The decision mainly relies on common knowledge which may not be suitable for the region (i.e., resulting from the unique environmental characteristics of the region). In this situation, there is a need for an efficient method to sample environmental data and to form an initial knowledge of the region. The data sampling process needs to be representative of the region and able to capture variations over location and time. This fourth research question will serve as a guide to construct an efficient data sampling technique to support the design of ESN.

1.3 Research Objectives

The work presented in this thesis aims to fill the gap in the current study of ESN design. The proposed method will consider redundancy as a factor to be balanced with robustness to form an optimum ESN design. The following comprises the main research objectives that will be discussed throughout this thesis:

- **Formulating the measure of representativeness.**
  An ESN is deployed in an RoI with the purpose of capturing certain environmental properties in that region. The ESN is expected to produce a set of measured data which best represents the region. Since representation is the main objective of the deployment, it has to be quantifiable. In this case, a clear formulation of representativeness measurement is needed. Such a measurement would enable the comparison of a particular ESN design against other ESN designs. This study aims to propose generic method to measure representativeness of an ESN.

- **Formulating ESN design as an optimisation problem.**
  Finding an optimal placement of sensor nodes within an RoI, given a number of nodes, is not a trivial task. Each set of node placements will yield a certain level of representativeness. In this case, the increase in either the number of sensor nodes or the size of an RoI would significantly increase the number of possible node placements in the region (i.e., search space). An optimisation technique which is able to deal with such a large search space is needed. This study aims to construct a method which optimises the sensor node placement, for a given number of nodes, which best represents the RoI.

- **Formulating data quality and robustness in ESN design.**
  Common ESN data quality issues are required to be identified and studied. They have to be clearly defined and well formulated, allowing their impact to the representativeness of an ESN design to be quantified. Some common techniques in overcoming the data quality issues (i.e., data quality controls) are also worthwhile to be explored and studied. These techniques, and the effectiveness of such techniques, should be included when considering ESN design. This study aims to include data quality issues and robustness considerations in the design of ESN.

- **Finding a balance between redundancy and robustness in ESN design.**
  In contrast to the common practice in ESN design, in this study, redundancy is not considered as a factor to be eliminated. Instead, redundancy is considered as a factor

to be balanced with robustness. In order to achieve this goal, both redundancy and robustness have to be clearly defined and formulated. Appropriate formulation of redundancy and robustness would enable the trade-off between these two parameters to be quantified. An optimisation technique which could optimise two factors is needed in order to explore all the possible ESN designs. This study aims to propose a method to balance the redundancy and robustness in the design of ESN.

- **Formulating an efficient data sampling technique to support the design of ESN.** Historical data related to the past measurement of environmental properties in an RoI is essential not only for the design process prior to the deployment of new ESN, but also for improvement of an existing ESN. Forming an ESN design without any access to historical data has never been a trivial task. It requires a substitution dataset to compensate for the absence of the past measurement data. In this case, a technique to build an initial knowledge regarding the environmental phenomena in the RoI is needed. As an extension, this study aims to propose a method which incorporates mobile data sampling to construct a baseline dataset of an RoI which will be utilised in the ESN design optimisation process.

1.4 Thesis Structure

The rest of this thesis is presented according to the following structure:

Chapter: 2 Literature Review Review of key literature to form a solid foundation for the experimental study being conducted in this thesis. Some key research areas are covered in this chapter including interpolation techniques, optimisation techniques, environmental sensor network, data quality, and design of environmental sensor network.

Chapter 3: Methodology Addresses problem formulation in ESN design within this study and discusses the proposed method to address the problem. The chapter starts with a description of the dataset being used in this study followed by the problem formulation in ESN design. The chapter then progresses with the description of the applied interpolation and optimisation techniques which are specifically chosen and tailored for the study. The design process is divided into two categories: with the knowledge of historical data and where the historical data is absent. Each component is formalised in clear and consistent mathematical notations/equations in order to improve the readability of this thesis. In addition, some diagrams/illustrations are provided to assist the reader's comprehension of the proposed method.

Chapter 4: Results and Validations describes how the proposed method (i.e., as described in Chapter 3) is applied in the form of experimental study. Results from each experiment are presented and findings are highlighted in this chapter.

Chapter 5: Discussions and Conclusions discusses and concludes several key components in this thesis which contribute to the body of knowledge in the area of environmental sensor network design. Some fundamental limitations of the study are also described in this chapter. Further, this chapter also highlights areas of ESN design, which may warrant further investigation into the future.

Chapter 2

Literature Review

This chapter review some key literature which form a firmed foundation for the experimental study being conducted in this thesis. Some key research areas are covered in this chapter including interpolation techniques, optimisation techniques, Environmental Sensor Network (ESN), data quality, and the design of ESN. The review of the literature is structured as follow:

- Section 2.1 Interpolation Technique
  This section describes fundamental concepts in data interpolation, some known interpolation techniques are also reviewed.
- Section 2.2 Optimisation Technique
  Core knowledge in optimisation problem is reviewed in this section, including the general formulation of decision variables and constraints (described in Section 2.2.1 and 2.2.2) which define a search space. This section also covers the formulation of objective function including the concept of domination and Pareto optimal (reviewed in Section 2.2.3, 2.2.4, and 2.2.5) where more than one one objectives are applied.
- Section 2.3 Environmental Sensor Networks (ESN)
  This section describes a brief background of ESN including the current development of ESN, the architecture of ESN, and also some applications of ESN (described in Section 2.3.1, 2.3.2, and 2.3.3 respectively).
- Section 2.4 ESN Data Quality
  This section describes some basic concepts of data quality in ESN and how it is essential. Common data quality issues in ESN are also reviewed, including some strategies to mitigate such issues which are categories in Quality Assurance and Quality Control (described in Section 2.4.1 and 2.4.2)

- Section 2.5 ESN Design
  This section describes how is design matter for having a fit for purpose ESN including the complexity in ESN design. This section also covers some challenges in ESN design, deployment strategy in ESN, and common objectives in the deployment of ESN (describes in Section 2.5.1, 2.5.2, and 2.5.3 respectively).

2.1 Interpolation Techniques

The need of spatially (and also temporally) continuous data of environmental properties are increasing in the environmental sciences. This information is not always readily available, and providing data related to the environmental parameters for any place at any time is a very challenging work for environmental scientists. Ideally, in order to achieve this goal, a considerably dense and interconnected sensor nodes are required to be deployed in the Region of Interest (RoI). This network would enable a fairly accurate estimation of the spatial distribution of the required environmental parameters. However, a network with a high dense sensor nodes is difficult and expensive to deploy and to maintain. Therefore, in most cases, the environmental parameters are measured limited at point locations only, sparse, and not on a regular grid. This often lead to a situation where the data is not available where it is most needed. Developing some methods to estimate the parameters in un-sampled locations is essential to overcome this limitation [1, 15, 101].

Interpolations techniques could be utilised as an alternative to estimate the spatial distribution of the climate parameters based on the measurements from the neighbouring sensor nodes. These interpolation techniques are also known as spatial interpolation methods. Formally, a spatial interpolation technique can be formulised as a mathematical function which capable to predict values at locations in space where there are no measured values available [45]. The uncertainties on the estimated values would increase considerably as the network density decreases. The spatial interpolation techniques are not only been utilised exclusively in environmental science, but also have been utilised widely in many other disciplines. There are several different techniques for spatial data interpolation are available. According to Li and Heap [100] Inverse Distance Weighting (IDW) and Ordinary Kriging (OK) are the most frequently used methods in environmental sciences.

2.1.1 Inverse Distance Weighting (IDW)

Inverse Distance Weighting (IDW) is a widely applied deterministic technique to interpolate spatial data. The estimated value is calculated based on a linear combination of values at sampled points weighted by an inverse of the distance from the point of interest (i.e., un-sampled point) to the sampled points (i.e., measured points). This method relies on the assumption that the nearby sampled points point have more similar values than the ones which are further away from the point of interest [52, 132].

The generic formula for IDW can be expressed as follows:

$$\hat{Z}(s_0) = \sum_{i=1}^{n} \lambda_i Z(s_i) \tag{2.1}$$

Where:

$\hat{Z}(s_0)$	is the estimated value located in s_0 (i.e., un-sampled point).
$Z(s_i)$	is the measured value located in s_i (i.e., sampled points).
λ_i	is the weight assigned for $Z(s_i)$ such that $\sum_{i=1}^{n} \lambda_i = 1$.
n	is the number of sampled points used for estimation.

The weights in IDW are calculated according to the following formula:

$$\lambda_i = d_i^{-p} \frac{1}{\sum_{i=1}^{n} d_i^{-p}} \tag{2.2}$$

Where:

λ_i	is the weight assigned for $Z(s_i)$
d_i	is the distance between s_0 and s_i (as illustrated in Figure 2.1)
p	is an exponent, also known as the power parameter
n	is the number of sampled points included in the interpolation

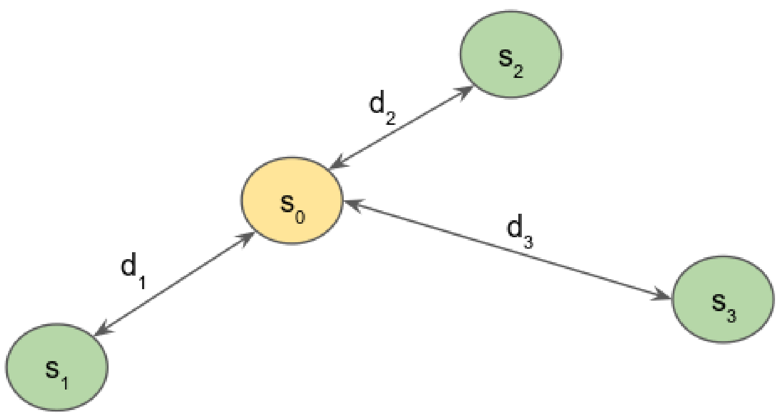


Fig. 2.1 Three sampled locations (e.g., s_1, s_2, s_3) are going to be used to estimate the value holds in un-sampled location s_0. The distance between the un-sampled location and each of the sampled locations s_1, s_2, s_3 are depicted as d_1, d_2, d_3 respectively.

Equation 2.2 suggests that the weights diminish as the distance increases, resulting a higher weight for the nearby sampled points which eventually bring more influence to the estimated value. In this case, the the power parameter p would greatly affect the accuracy of the prediction. The selection of the power parameter p and sample size n is arbitrary. Two is the most commonly used value for p, which makes IDW also known as the Inverse Distance Squared (IDS) [115]. Depending on what value been assigned to p, IDW also addressed as "moving average" in the case of p is zero, "linear interpolation" when p is 1 and "weighted moving average" when p is not equal to 1 [99]. The relationship between weight, distance, and power parameter in IDW is presented in Figure 2.2

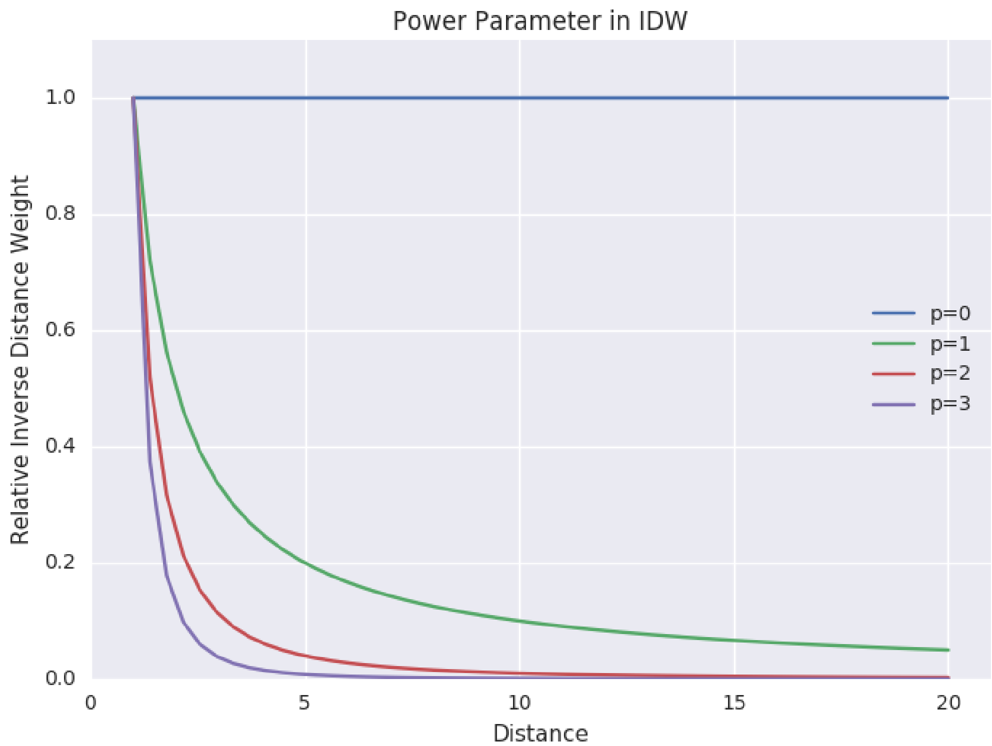


Fig. 2.2 The relationship between the power parameter p and the inverse distance as a weight in interpolation method, as employed in IDW. Adjusting p to zero means no weights will be implemented and all sample points will be treated equally. As presented in the figure, as p increases, the influence of the nearest sample points will increase and reduce the influence of the farther ones; resulting a more detail interpolated surface. On the contrary, reducing the power parameter p will allow more influence of the farther sample points resulting a smoother interpolated surface.

2.1.2 Ordinary Kriging (OK)

Ordinary Kriging (OK) is one of the geostatistical methods to estimate the value in an unsampled location based on the measured values from nearby sampled locations [38]. Similar to IDW, OK also implements weights in its calculation. The generic formula of OK is also similar to IDW as described in Equation 2.1. In contrast to IDW, OK assigns weights to its sample points not only based on their distances, but also based on the spatial variability structure. In OK, spatial and statistical relationships are considered as the basis to constructs weights [151].

Semivariance is employed to measure the the degree of spatial dependence between sample locations. In terms of calculation, semivariance is simply a half of the variance of all available sample points in space with a constant distance apart. In geostatistics, the semivariance is formalised as follows:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (Z(s_i) - Z(s_i + h))^2 \tag{2.3}$$

Where:

$\gamma(h)$	is the semivariance with lag distance h.
h	is the lag distance between two sample points.
$N(h)$	is the number of sample points which have lag distance of h with the other sample points.
$Z(s_i)$	is the value measured in sample point s_i.

The semivariance can be estimated from the data by employing certain fitted function (i.e., variogram modelling) and the plotting of the fitted data (i.e., variogram). This modelling and estimation is essential for structural analysis and spatial interpolation [24]. Figure 2.3 shows an example of variogram plotting.

The weights in OK are obtained by minimising the variance in OK prediction error $[\hat{Z}(s_0) - Z(s_0)]$, also known as "Kriging variance", which is formalised as follows:

$$\sigma_e^2 = \sum_{i=1}^{n} \lambda_i \gamma(x_i, x_0) + \theta \tag{2.4}$$

Where:

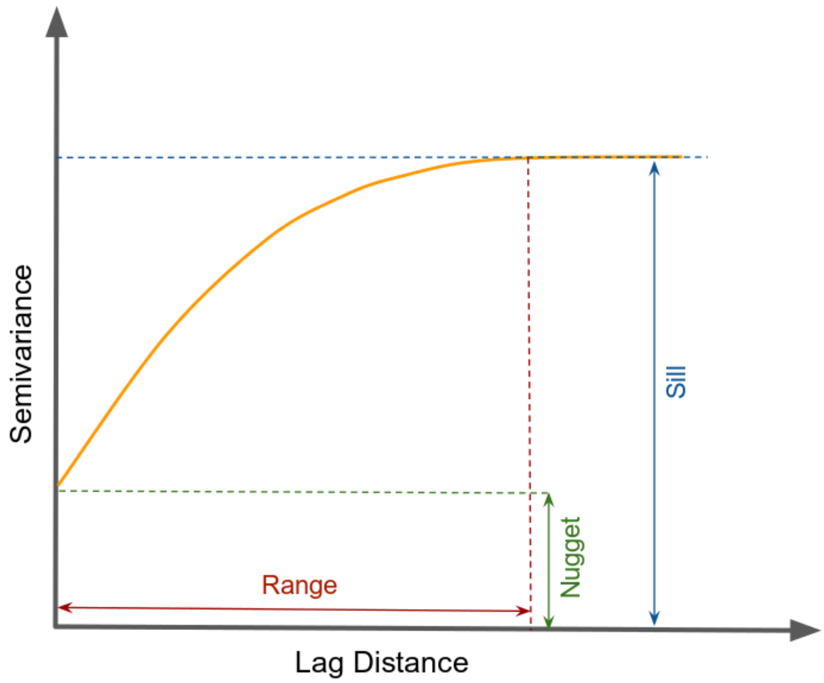


Fig. 2.3 An example of a variogram with range, nugget, and sill. In variogram, the “range” indicates the shortest distance at which the “sill” is reached. The “range” could be used to identify the size of a search window used in the interpolation methods, where samples with distance larger than the range are spatially independent and not included in the interpolation process. A positive value of the semivariance at lag distance close to zero is called the “nugget”, which also indicates the variance of sampling errors and the spatial variance at shorter distance than the minimum sample spacing [99].

σ_e^2	is the Kriging variance.
n	is the number of sample points.
λ_i	is the proposed weight for sample point s_i such that $\sum_{i=1}^{n} \lambda_i = 1$.
$\gamma(x_i, x_0)$	is the semivariance between the values at sample location s_i and un-sampled location s_0. Such value can be obtained from the fitted variogram.
θ	is the Lagrange multiplier required for minimisation.

2.2 Optimisation Techniques

Optimisation can be defined as the task of finding one or more feasible solutions, which can produce desired or useful values of one or more objectives. Optimisation itself can be a single-objective optimisation or multi-objective optimisation. If there is only one objective function required to be satisfied, the task of finding the optimal solution is called single-objective optimisation. It follows then that if there is more than one objective function that needs to be satisfied, the task of finding one or more optimum solutions is called multi-objective optimisation [22, 42]. Further, Coello et al. [35] provides a clear definition of multi-objective optimisation problem as:

> "The problem of finding a vector of decision variable values which satisfies constraints and optimises a vector function whose elements represent the objective functions. These functions form a mathematical description of performance criteria, which are usually in conflict with each other. Hence, the term 'optimise' means finding such a solution which would give the values of all the objective functions acceptable to the decision maker."

Most problems in real world applications have multiple objectives, which are possibly conflicting with each other. By optimising one objective, one may be sacrificing the other objectives. A simple example can be found in computing equipment purchase decisions. People in general want to have computing equipment with high performance. However, people also want to save their money and spend less. In this case, the objective of having the computing equipment with the best performance cannot be achieved without abandoning the objective of spending less money in purchasing. On the other hand, the objective of spending less cannot be achieved without sacrificing the objective of having computer equipment with the best performance. The objectives in the purchasing decision are in conflict with each other.

2.2.1 Decision Variables

The decision variables in optimisation problems are the numerical values, which are chosen in such a problem. In mathematical notation, the variables can be represented as:

$$x_i, i \in \{1, \cdots, n\} \tag{2.5}$$

Or it can also be noted as a vector x of n decision variables as follow:

$$\bar{x} = [x_i, \cdots, x_n] \tag{2.6}$$

A vector of n decision variables in an optimisation problem is also known as a solution.

2.2.2 Constraints and Decision Variable Bounds

Constraints in optimisation problems are the restrictions or limitations introduced by the environment or resources, such as physical limitations, time restrictions, processing power limitations, and several other kind of limitations. Certain solutions can be considered acceptable when these solutions can satisfy all the available constraints. In mathematical notation, these constraints can be represented either in mathematical inequality:

$$g_i(\bar{x}) \leq 0, i \in \{1, \cdots k\} \tag{2.7}$$

or equality as follows:

$$h_j(\bar{x}) = 0, j \in \{1, \cdots l\} \tag{2.8}$$

In this case, k represents the number of inequality constraints and l represents the number of equality constraints.

Let n be the number of decision variables, then the number of inequality constraints k cannot be greater than or equal to n. In other word, k must be less than n ($k < n$). Since the degree of freedom in multi-objective optimisation problem is defined as $n - p$, therefore, the optimisation problem with $k \geq n$ is considered as over constrained and there is no more flexibility or any degree of freedom for optimising.

In addition to the constraints, there are also decision variable bounds. In mathematical notation, the variable bounds can be expressed as follows:

$$x_i^{(L)} \leq x_i \leq x_i^{(U)}, i \in \{1, \cdots, n\} \tag{2.9}$$

Where:

$x_i^{(L)}$ is the lower bound value for decision variable i.

$x_i^{(U)}$ is the upper bound value for decision variable i.

n is number of decision variables to form a solution.

These variable bounds restrict each decision variable to take a value only in the range between the lower value and the upper value. The bounds also represent a decision variable space D known as decision space.

A solution $\bar{x}$ is defined as a feasible solution if and if only it satisfies all of the constraints and the variable bounds. On the other hand, if any solution $\bar{x}$ does not satisfy all the constraints and the variable bounds, it is known as an infeasible solution. Clearly, not all solutions in the entire decision variable space D are feasible solutions. The set of all feasible solutions is known as feasible region S.

2.2.3 Objective Function

In the study of optimisation problems, an objective function is defined as the computable function of a vector of decision variables, which is used as a criterion to evaluate a certain solution in order to know how good the solution is. In real world optimisation problems, some functions are required to be minimised while other functions are required to be maximised. Moreover, in multi-objective optimisation problems, these functions in many cases are in conflict with each other. Optimising a particular objective function may sacrifice the other objective functions. These objective functions may be measured using the same measurement units (i.e., commensurable) or the functions may also be measured using different measurement units (i.e., non-commensurable). In mathematical notation, the objective functions can be represented as follows:

$$f_i(\bar{x}), i \in \{1, \cdots, m\} \tag{2.10}$$

Where:

$f(\bar{x})$	is the objective function to be optimised given $\bar{x}$ as the parameter.
$\bar{x}$	is a vector of decision variables, also known as a solution.
m	is the number of objective functions being solved in the multi-objective optimisation problem.

Since there would be more than one objective function available in a multi-objective optimisation problem, these functions will form a vector function, which can be expressed in mathematical notation as follows:

$$\overline{f(\bar{x})} = [f_1(\bar{x}), \cdots, f_m(\bar{x})] \tag{2.11}$$

Referring to the notation, the goal in a multi-objective optimisation problem can clearly be seen as the problem of optimising the m number of objective functions simultaneously. The optimisation process itself can be maximising the values of all m objective functions, or minimising the values of all m objective functions, or even in some cases could be combining the maximisation and the minimisation values of these m objective functions. Since the task in multi-objective optimisation problems is about optimising a vector of objectives instead of a single-objective, multi-objective optimisation is also known as vector optimisation.

Most of the optimisation algorithms that have been developed deal with only one type of optimisation problem, which is either the problem of minimising or maximising. In order to simplify the task of dealing with mixed types of optimisation problems, the duality principle can be applied. In the context of optimisation, the duality principal suggests that a maximisation problem can be converted into minimisation problem by multiplying the objective function by negative one (-1). The same thing works vice versa, depends on the implementation of the algorithm.

2.2.4 Concept of Domination

The concept of domination is widely applied in the field of multi-objective optimisation problems in order to compare two solutions. Two solutions are compared to see whether one solution dominates the other or not.

Let $\bar{x}$ and $\bar{y}$ be two solutions in a multi-objective optimisation problem. Solution $\bar{x}$ is said to dominate solution $\bar{y}$, or in mathematical notation expressed as $\bar{x} \preccurlyeq \bar{y}$, if and if only it complies with these two domination conditions:

1. Solution $\bar{x}$ is no worse than solution $\bar{y}$ in all objective functions;
2. Solution $\bar{x}$ is strictly better than solution $\bar{y}$ in at least one objective function.

In the case of minimisation as an optimisation problem, these domination conditions can be expressed in mathematical notation as follows:

$$\begin{aligned} \bar{x} &= [x_1, \cdots, x_n] \\ \bar{y} &= [y_1, \cdots, y_n] \\ \bar{x} \preccurlyeq \bar{y} &\Leftrightarrow (\forall i : f_i(\bar{x}) \leq f_i(\bar{y})) \wedge (\exists i : f_i(\bar{x}) < f_i(\bar{y}), i \in \{1, \cdots, m\}) \end{aligned} \tag{2.12}$$

Where n represents the number of decision variables that construct a solution and m represents the number of objective functions being solved in a multi-objective optimisation problem.

Apart from representing solution $\bar{x}$ dominating solution $\bar{y}$, this mathematical notation also implies that:

1. Solution $\bar{y}$ is dominated by solution $\bar{x}$,
2. Solution $\bar{x}$ is non dominated by solution $\bar{y}$,
3. Solution $\bar{x}$ is non inferior to solution $\bar{y}$.

2.2.5 Pareto Optimal

In a single-objective optimisation problem the notion of optimality can be clearly identified. The optimum solution can be found by simply looking for the best value of the predefined objective function. On the other hand, in a multi-objective optimisation problem there is more than just a single objective function require to be satisfied and in most cases, the objective functions are in conflict with each other. Finding a single global optimal solution in the decision variable space D is nearly impossible. In multi-objective optimisation problems, instead of looking for a single solution, the focus is looking for a trade-off among the objective functions. For this purpose, optimality of a solution needs to be redefined properly in order to respect the integrity of each objective function [34, 43, 64].

The concept of domination is utilised in the multi-objective optimisation problem and it is also known as the Pareto dominance. All possible pairwise comparisons can be performed for a given finite set of solutions in order to find which solutions are non-dominated with respect to each other. The set of non-dominated solutions that is left has the property of dominating all other solutions apart from the solutions which belong to this set. Any member in the entire search space of solutions does not dominate these solutions. In other words, the set of non-dominated solutions are better compared to all other solutions [79, 147].

In multi-objective optimisation problem, the set of non-dominated solutions is also known as Pareto Optimal Set. In mathematical notation, the Pareto Optimal Set can be expressed as follow:

$$P^* = \{\bar{x} \in D | \neg \exists \bar{x}^* \in D : f(\bar{x}^*) \preccurlyeq f(\bar{x})\} \tag{2.13}$$

Where:

P^*	is the set of non-dominated solutions, also known as Pareto optimal set.
$\bar{x}$	is a solution (i.e., a vector of decision variables).
$\bar{x}^*$	is an optimal solution.
D	is the search space where a solution could be found/formed.
$f(\bar{x})$	is the objective function to be optimised given $\bar{x}$ as the parameter.

The global Pareto Optimal Set can be defined as the non-dominated set of the entire feasible search space S . Often the globally Pareto Optimal Set is simply referred to as Pareto Optimal Set.

Furthermore, by plotting the Pareto Optimal Set in objective space, the non-dominated vectors are collectively known as the Pareto Front. Figure 2.4 shows an example of Pareto Front with two objective functions. In mathematical notation, Pareto Front can be represented as follow:

$$PF^* = \{u = f(\bar{x}) | \bar{x} \in P^*\} \tag{2.14}$$

Where:

PF^*	is the Pareto Front.
P^*	is the set of non-dominated solutions, also known as Pareto optimal set.
$\bar{x}$	is a solution (i.e., a vector of decision variables).
u	is a value produced by the objective function f given a vector of decision variables $\bar{x}$ as the parameter.

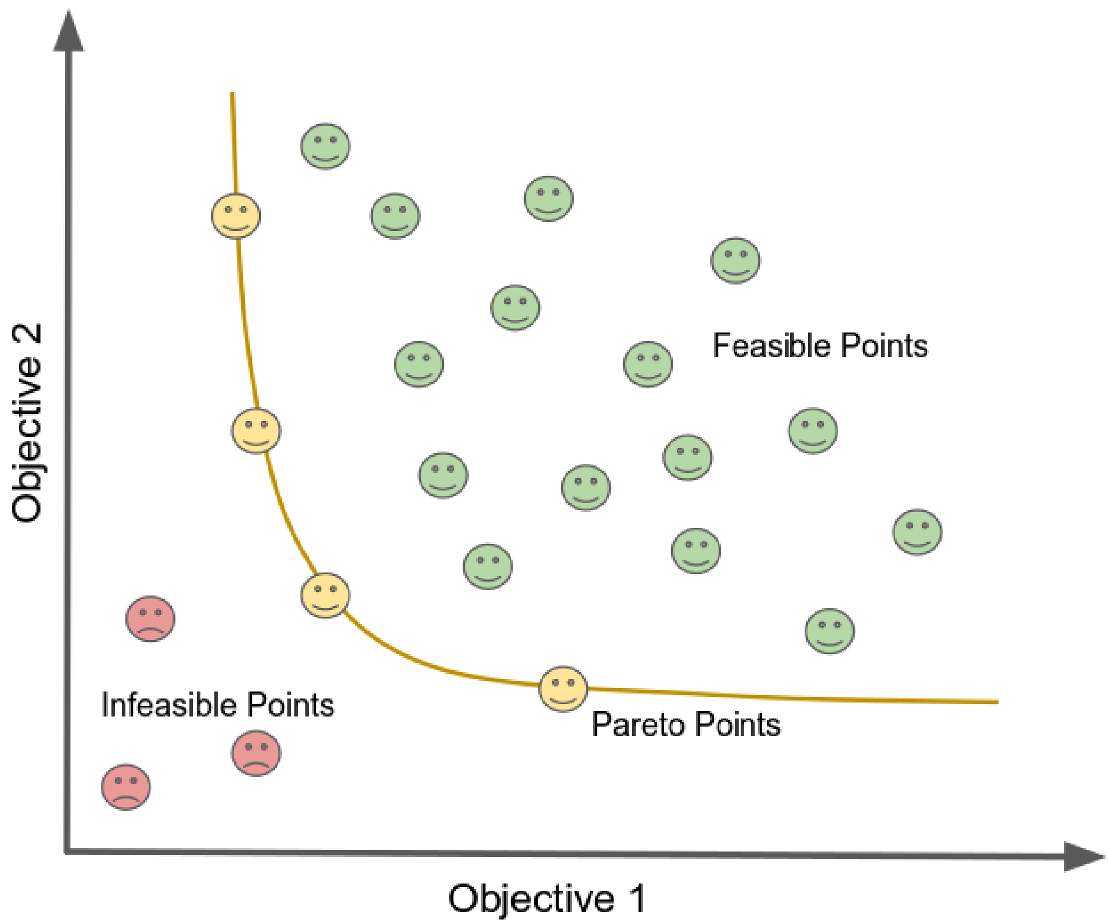


Fig. 2.4 The plotting of solutions in objective space, where there are two objective functions to be optimised. Infeasible points represent the points in the objective space which violate the constraints (as described in Sub-section 2.2.2). The points in objective space which obey the constraints are called the the feasible points, which also includes the Pareto points. The Pareto Points are the feasible points in objective space which dominate other feasible points but not dominated by the others. The Pareto points also known as the non-dominated points, which collectively will form the Pareto Font (depicted as the yellow line in the figure).

2.2.6 Evolutionary Algorithms

In order to solve multi-objective optimisation problems, the Operations Research community has developed several approaches since the 1950s based on a variety of mathematical programming techniques. However, there are several limitations in mathematical programming techniques when dealing with multi-objective optimisation problems. Most of them only produce a single solution for each run; therefore in order to produce a Pareto Optimal Set, several runs are required. Moreover, mathematical programming techniques in general are susceptible to the shape and continuity of the Pareto Front [35, 57].

Evolutionary Algorithms are computer programs that mimic natural evolutionary principles, which are inspired by Charles Darwin, in order to solve complex searching and optimisation problems. In Evolutionary Algorithms there would be a number of artificial creatures, known as individuals, which are generated to search over a particular problem space. Individuals continually compete against each other in order to discover the optimal areas from the predefined search space. Gradually, over some periods of time, the most successful individuals evolve to discover the optimal solution [28, 43, 64].

In contrast to the mathematical programming techniques, which in general only produce a single solution for each run, Evolutionary Algorithms can find several members of the Pareto Optimal Set in a single run. Evolutionary Algorithms are also less susceptible to the shape or continuity of the Pareto Front.

The individuals in Evolutionary Algorithms are commonly represented by strings or vectors that have a fixed length. Every individual encodes a unique possible solution to address a particular problem. In Evolutionary Algorithms, a set of individuals is known as a population.

The Evolutionary Algorithm is started with an initial population consisting of a particular number of randomly generated individuals. A fitness value is then calculated for each individual. In order to generate the fitness value, each individual is decoded to produce a possible solution to the problem. The fitness function will calculate the solution value to produce a fitness value for the corresponding individual. The individuals with higher fitness values represent better solutions to address the problem, compared to the ones with lower fitness values. This initial process is followed by the main iterative cycle, which consists of two main operations, mutation and recombination [35]. Figure 2.5 presents the recombination and mutation process which is commonly used in Evolutionary Algorithms to maintain variation within the population.

For each iteration, the individuals in the current population produce a new set of individuals called children (i.e., offspring). After the fitness value is assigned to every child, a new population is created. The current individuals and the children are allocated to become

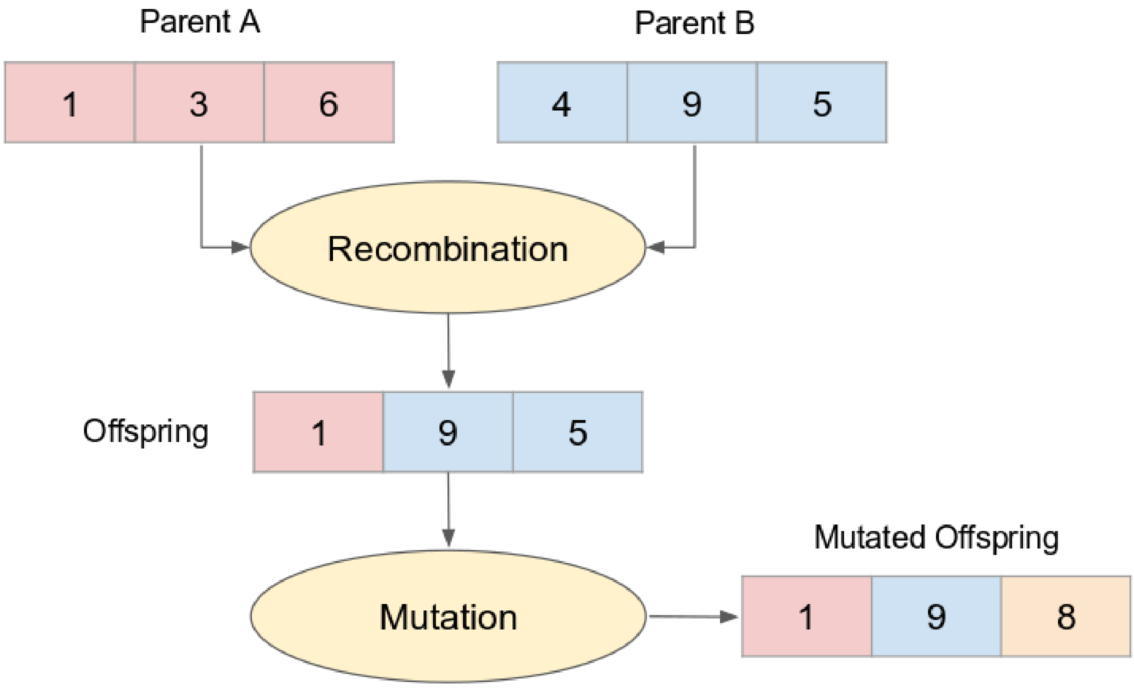


Fig. 2.5 The recombination process will exchange the features or characteristics between parents to form a new individual, also known as offspring. The mutation process will randomly alter certain features/characteristics of an individual. Recombination and mutation are utilised in Evolutionary Algorithms to maintain variation within the population.

members of the new population. This new population will be treated as the current population in the next iteration cycle. In order to control the growth of the population, a similar approach to the natural evolutionary strategy (the survival of the fittest) is applied and the individuals start competing against each other. This kind of approach in Evolutionary Algorithms is known as the selection process. The fitness value is used as the basis for the selection process. The individuals with better fitness values have more chance of being selected as parents (to produce offspring) and also to be selected to form a new population. Such iterative process will run until certain termination condition is satisfied. Maximum number of generation is a commonly used termination condition in Evolutionary Algorithms. Figure 2.6 shows the overall process in Evolutionary Algorithms.

According to Deb [42], in order to solve multi-objective optimisation problems, there are four main primary goals that can be identified in Evolutionary Algorithms:

1. Maintain the non-dominated points in the objective space and associated solution points in the decision space.

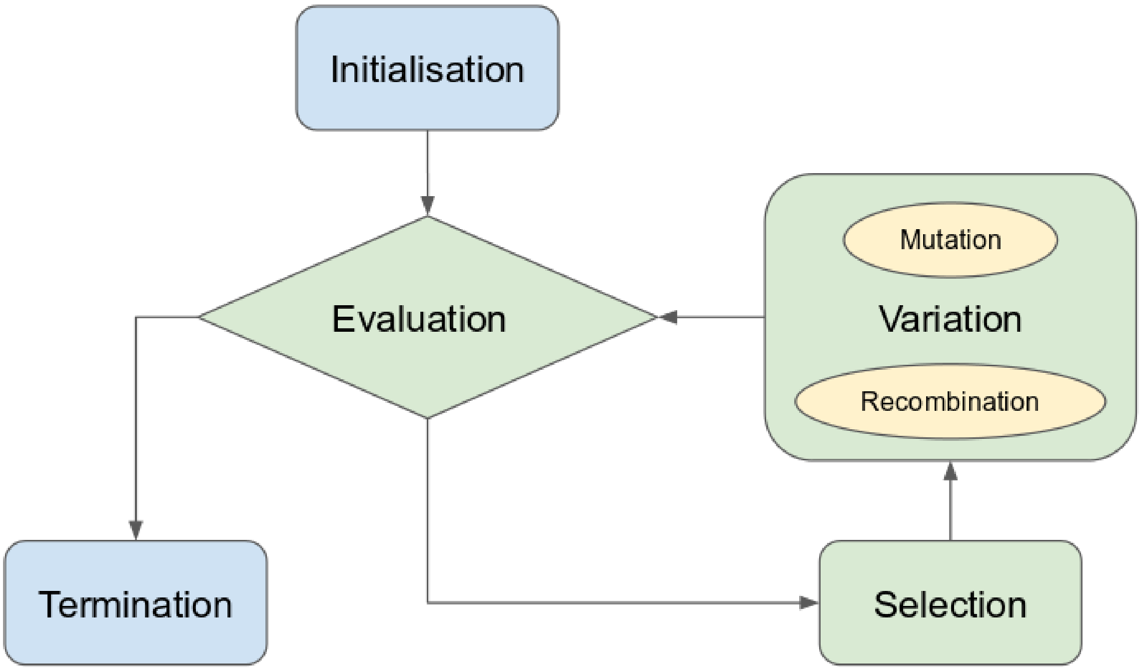


Fig. 2.6 A general work-flow in an Evolutionary Algorithm. The initialisation part will generate number of random individuals to form an initial population. The evaluation part takes care two main tasks: evaluate the termination condition and evaluate each individual within the current population in respect to all fitness functions. The selection part in-charge in forming a new population for the next generation. The individuals with higher fitness values will be chosen to form the new population. Variation within the population is essential in order to explore the search space and to avoid the local optima. In Evolutionary Algorithm, the variation is maintained through mutation and recombination process.

2. Continually make algorithmic progress towards the Pareto Front in the objective function space.
3. Maintain diversity in the Pareto Front Set and the Pareto Optimal Set.
4. Provide a large enough Pareto Optimal Set for the decision maker.

2.3 Environmental Sensor Network (ESN)

Automated environmental monitoring started with simple automatic logging systems that continuously recorded several environmental properties at predetermined intervals. These simple monitoring systems had no communication capability. They required field scientists to visit the site regularly and download the data manually. Technological advancements enabled these passive logging systems to evolve into intelligent sensor networks where each sensor node actively communicates its own observed data to nearby sensor nodes. Moreover, these interconnected sensor nodes also have a capability to process and communicate their data to a remote data centre without any operator intervention. These monitoring systems are known as Environmental Sensor Networks (ESNs), which enable long-term environment monitoring at scales and resolutions that are difficult to achieve with conventional observation methods [32, 37, 68, 110, 121].

2.3.1 Development of Sensor Networks

The deployment of cheap and smart devices in large numbers with multiple on-board sensors, which connect together through wireless networks and the Internet, offers tremendous and unprecedented opportunities for collecting information on a wide range of entities of interest. Even though the research on sensor networks was initiated for military purposes, further development in low-cost sensors and communication networks has broadened the potential application of sensor networks from infrastructure security to industrial sensing. The sensor networks could be deployed in houses, offices, hospitals, cities, and the environment to get a better understanding and to control surrounding conditions [32].

Three different areas of study are involved in the development of sensor networks technology: sensing, communication, and computing, which includes hardware, software, and algorithms. Development in sensor networks has been driven by both the combined and separate advancement in each of these three research areas. The collaboration can clearly be seen from the three main components of a sensor node: sensors as sensing devices, data processing unit, and communication unit which enable it to establish untethered communication at short distance [4, 32, 92].

An on-board processing unit enables each sensor node in sensor networks to conduct in-situ data processing. Therefore, rather than sending the raw data to the base station, each node can handle simple data processing locally and transmit only the required data to the base station. This cooperative effort of sensor nodes is one of the unique properties of sensor networks [3, 121]. Furthermore, the advancement in Micro-Electro-Mechanical Systems (MEMS) brings a substantial contribution in the miniaturisation of sensor nodes. MEMS has revealed the possibility of producing significantly smaller sensor nodes with multifunctional capability and low power consumption. The fast growth in the development of MEMS offers a wide range of promising future applications, considering the relatively low manufacturing cost offered by this technology [69].

Wireless communication in sensor networks is unique in terms of features and requirements, which are different compared to traditional wireless ad-hoc networks. The communication protocol in sensor networks has to deal with limitations in power, processing capability, and memory capacity. In a sensor network, sensor nodes are densely deployed in large numbers and the network topology changes frequently. The individual nodes are also prone to failure. Considering the large number of sensors and the communication overhead, having a global identification for sensor nodes, which is commonly found in traditional networks, may not be possible. Broadcast communication is also employed by most sensor nodes, whereas ad-hoc networks mainly rely on point-to-point communication. These conditions make the technologies that are currently available for ad hoc networks not well suited to the unique requirements of sensor networks. Recently, there have been many researchers working in this area in order to fulfil these requirements [4].

Energy has been recognised as one of the main challenges in ESN, especially for those deployed in remote locations. The recent development in energy harvesting technology could be considered an important element supporting the advancement in ESN. This particular technology enables electronic devices to produce some amount of energy (electrical power) from its ambient environment such as by utilising propagated radio waves, wind flow, sunlight, or mechanical vibration [12].

2.3.2 Sensor Network Architecture

A generic ESN architecture is constructed with three main component: sensor nodes, base stations, and a sensor network server (as presented in Figure 2.7). Sensor nodes gather the environmental raw data and simple data processing might be present in each node. The pre-processed data will be passed to one or more base stations for further data processing. Sensor Network Server (SNS) acts as a data repository where data from several base stations is aggregated. More sophisticated data processing will also be handled at this stage. In order to provide seamless access to the environmental information for external users, a web service is utilised as an interface between the SNS and the users. Moving up the hierarchy from the sensor nodes to the SNS, there is an increase in computational capability, data storage capacity, and power availability. In general, the sensors nodes and the base stations in ESN may be able to serve for a few months only. This is due to the power supply limitation and harsh environmental conditions [4, 3, 5, 136].

2.3.3 Applications of ESN

ESN has a significant role to support the quality of our life on this planet. They play a part in many different areas such as agriculture, forestry, science, healthcare and safety, insurance, mining, weather forecast, etc [8, 113, 156].

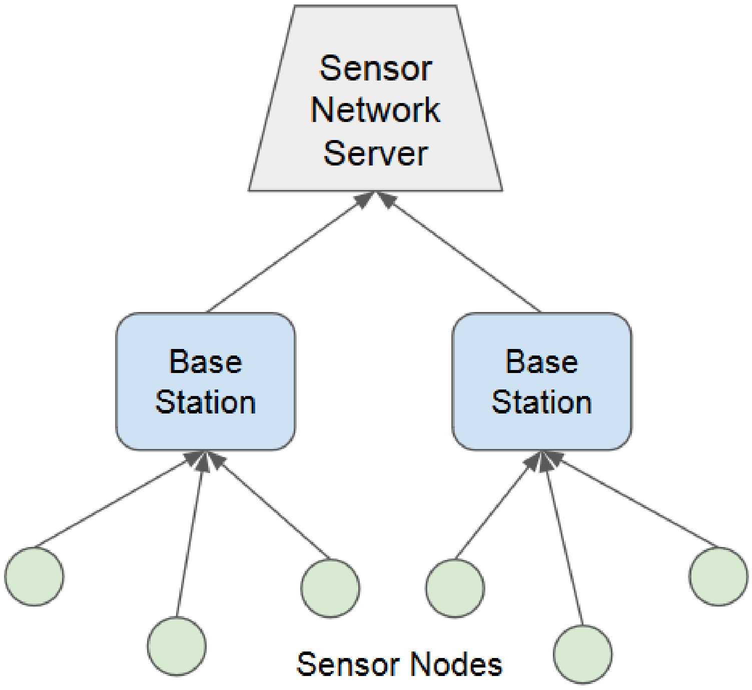


Fig. 2.7 A generic ESN architecture which consist of three main components: sensor nodes, base stations, and a server known as Sensor Network Server (SNS). Environmental parameters in the Region of Interest (RoI) are measured and recorded by the sensor nodes. The data is then passed to one or more base stations.

Production in agriculture and forestry are highly dependent on the changes in the environmental parameters (e.g., temperature, humidity, rain fall, solar radiation). Apart from the need to increase production, agricultural management should also be practiced with a degree of precision (i.e., Precision Agriculture) to provide an alternative and realistic means to reduce the use of potentially harmful compounds and promote sustainability. Precision Agriculture is an emerging area where ESN plays an important role [14, 71, 93, 118, 120].

In forestry, ESN is also utilised for fire detection systems [108, 112, 144]. The networks can alarm the origin of the fire before it is spread uncontrollably. A major forest fire can destroy thousands of hectares and incur social, environmental and economic costs.

As living beings, water and air are crucial to support our life. The quality of the water and the air which we consume and breath every day directly impact our health. The need to promote better health-care also motivate the extensive use of ESN to monitor the quality of both water [47, 86, 137] and air [51, 109, 133].

In the scientific field, ESN enables us to have a better understanding of the planet on which we live. It helps us to answer many questions which could not be answered in the past and also to promote more questions which have never been asked before. The changes in climate across the earth would never be able to be identified without ESN. Nowadays, scientists around the world have more data than before to unveil the climate change and deeply analyse its impact [62, 78].

The advancements in ESNs also benefit our day to day life by providing more accurate weather information (e.g., weather report and forecast), which are crucial in certain areas like tourism and transportation. Such information are also be used to support personal decisions as simple as deciding what kind of clothes to wear in the day to suit the weather.

2.4 ESN Data Quality

Apart from the significant contribution brought by sensor networks to environmental monitoring, they are susceptible to malfunctions that may lead to data lost or poor data quality. Typically, sensors do not produce data with good quality all the time; abnormality in sensor data is common and it should be anticipated and planned for. The issue can be related to hardware, software, communication links, or environmental conditions. Some natural phenomena such as floods, fire, lightning strikes, and animal activities could disturb the functionality of the sensors. Moreover, malicious human activities such as stealing and vandalism also contribute to sensor failure. Malfunction in sensors could easily happen especially when they are deployed in unsuitable environments or without proper maintenance [36, 123, 162].

Peppler et al. argue that the performance of ESN systems can be measured using three parameters: availability, usability, and accessibility of the data to a user. In order to fulfil the availability aspect, the ESN system must able to collect, process, and deliver the data to a central repository and make it available to the users in a timely manner. High performance ESN systems should also be able to produce useable data with sufficient quality for certain purposes. The systems also able to notify the user if there is any known anomaly in the data. The ESN data must be accessible; where the users can easily find and retrieve the data they need from the central repository, without experiencing any difficulty while working with the data.

Delivering raw environmental data produced by an ESN to end users is unfortunately common practice in the ecological community. The sensor data is delivered with limited or no quality control. This situation is mainly triggered by the sheer volume of the data and also the large variety of sensor types which make them challenging to manage. Considering the huge volumes of the data produced by sensor networks and the time constraints imposed by near real time data processing, manual methods for sensor data quality assurance (QA) and quality control (QC) are no longer adequate. There is a potential for erroneous or misleading results in utilising the data with the absence of comprehensive checks or evaluations. This situation has motivated the study of automated QA/QC in ESN [26].

The term QA and QC are often used together. Even though they are closely related, they have distinct meaning. In the context of data quality in ESN, Campbell et al. define QA as:

> A set of processes or steps taken to ensure that the sensor network and protocols are developed and adhered to in a way that minimises inaccuracies in the data produced.

And QC is defined as:

> A process to identify and flag suspect data after they have been generated.

The definition clearly describes QA as process oriented with the main objective to produce high quality data with minimum need for data corrective measures. In other words, QA is a proactive or preventive process to avoid problems that may lead to poor data quality. It takes part while data is being produced. In contrast to QA, QC is product oriented with the main objective to identify and flag suspect data of poor quality generated by the sensor. In other words, QC is an evaluation process that assesses whether the data produced by the sensor satisfy the requirements for quality specified by the end users. It takes part after the data is generated by the sensor.

2.4.1 Quality Assurance (QA)

The value of environmental monitoring relies on the accuracy and precision of the data, which it represents the physical properties being measured. A good or bad data set is mainly driven by several factors that influence the data collection process, such as instrument calibration, long-term field exposure to the elements, and instrument maintenance. Therefore, a comprehensive end-to-end QA procedure (from sensor node deployment to calibration and maintenance) is essential [119].

Sensors require regular maintenance and scheduled calibration in order to reduce data loss and produce high quality data. Calibration drift is a common anomaly in sensor data where sensor components deteriorate over time because of age related processes such as corrosion, fatigue, and photo degradation. The instrument calibration procedures include the calibration prior the deployment and periodic calibration during the operational phase, represent crucial components of the QA process. The calibration process may be as simple as side-by-side comparisons of measured data conducted during the routine maintenance visits or as complex as laboratory comparisons to known standards. Ideally, the calibration process is done based on manufacturer recommendations, however, some adjustment often required to suit the remote operation. In some cases, the calibration routine can only be done by the manufacturer [87].

As the second component of QA, maintenance process in ESN has a primary objective to ensure the performance and reliability of the instrument and the site. The process consists of a cycle of structured activities that result in a continuous and repeatable effort. As an example, in submerged sensors, calibration drift is often caused by bio-fouling where regular cleaning is required to control this issue. A reliable maintenance capability also requires timely procurement of parts and services to repair or replaced the failed components. Therefore, having some replacement parts on site is necessary to ensure that any part of the network can be replaced immediately in case of damage or destruction [26].

The deployment of redundant sensor nodes might also be involved in QA to minimise data loss due to sensor failure especially in the case where the data is crucially important. One good example is the ESN operated by The National Oceanic and Atmospheric Administration's US Climate Reference Network [46]. The network has been running for a decade with the main objective to measure and collect national meteorological data, which includes air temperature, precipitation, soil moisture, and soil temperature. In order to enforce the high data quality produced by the networks, they are deployed with triple redundancy across the 114 sites. By employing redundancy in the networks, a high degree of confidence in national climate data over the long term could be achieved.

2.4.2 Quality Control (QC)

QC is necessary to ensure that the collected data is fit for purpose. Ideally, the proper functioning QC process will only accept valid or good data and reject all erroneous or bad data. However, some false detection might occur in real world applications. The case where good data are falsely identified and marked as bad data is known as false positive. On the other hand, a false negative arises when erroneous data are accepted as good data [26, 142].

A very simplistic approach to identify any kind of abnormality in sensor data is by comparing it with the data produced by nearby sensor nodes. If a sensor node within a group of neighbouring sensor nodes produces a significantly different result from all the others, a further investigation of the node might be required since it could be the indication of abnormality in the sensing process [162].

Apart from minimising data loss due to sensor failure, redundancy in the deployment of sensor nodes can also beneficial for detecting anomalies in the data produced by the sensors. Subtle anomalies, such as calibration drift, are often difficult to detect without employing redundancy in sensor nodes. Three replicate sensor nodes is a typical minimum number required to detect the drift, since it is difficult to determine which sensor is drifting if there are only two replicate sensor nodes [26].

For small scale sensor data, manual methods in QC might be sufficient, however, considering the significant growth in the data produced by large scale ESN, these manual approaches are often no longer practical. As the volume of data being collected by sensor networks grows, automated QC procedures are becoming increasingly essential. Such automatic procedures will promote more accurate and faster identification of anomalies in sensor data, with the absence of human error [122].

Despite the urgent need of automatic QC procedures in ESN, these automated procedures are also known of their drawback from producing a significant amount of false positive detections. This is the condition where the valid observed data are mistakenly identified and flagged as erroneous data. In order to overcome this issue, some systems implement the QC procedure in a semiautomatic fashion. These systems would employ trained human operators to inspect the flagged data (i.e., data detected by the automatic procedures as erroneous data) and remove the flag whenever the operators deem the data to be valid. However, this semiautomatic procedure would not be sufficient for future needs. Therefore, the challenge is to deign an automated QC system without involving any human intervention hence effective at detecting a wide range of data errors with a low rate in false positive event [49].

Defective or missing data in ESN are inevitable and can adversely affect the value of the data, especially when the users who consume the data are not familiar with the measurement methods and conditions that may have caused the anomalies. In this situation, it would be very difficult for the users for recognising and correcting the errors [26].

In some degree, erroneous data are still usable under certain treatment. Ideally, data correction should be included in QA/QC procedures. The data correction procedures might cover correction of out of range values, correction for instrument fouling and drift, correction of anomalous values, and correction of any known bias in the sensor data. Any anomaly in the data requires decisions on whether to remove, adjust, or replace the data with an estimated value. However, data correction procedure is not trivial. It is a complex endeavour and can lead to misinterpretation and inappropriate data use. Moreover, the decision about whether to fill gaps in the data and the selection of the method with which to do so are subjective. Some factors such as the length of the gap, the level of confidence in the estimated value, and how the data are being used might be considered in order to support the decision making process [72].

2.5 Designing ESN and its Challenges

Design of an ESN is about deciding how many sensor nodes are required to best represent a given region, as well as where those nodes will be deployed, how frequently they should collect and communicate the data, and for how long they will operate. Design of ESN could also inform priorities for maintenance, the impact of sensor node failure, and even the quality of the sensors and their supporting hardware. Designing a sensor network is intrinsically an optimisation problem.

In order to have a fit for purpose ESN, design is a critical process prior to the deployment phase. There are two fundamental questions that need to be addressed: how many sensor nodes are required to fit the application purposes and where should the nodes be deployed in the Region of Interest (RoI) [44, 117].

In current ESN design practice, one of the major focus for reducing costs is to minimise the total number of sensor nodes required to cover a specific RoI [53]. However, when sensors fails, the usefulness of the network degrades. The ESN no longer produces the data needed; it is not advisable, or even possible, to rely on data from such a network for decision-making. Improving robustness of ESNs is paramount.

Complexity is introduced especially when dealing with the requirement to have a fully operational ESN, which meets the application purposes, with the lowest possible number of sensor nodes. Moreover, the uncertainty in a sensor's ability to function properly, resulting from disruptions that may be caused by terrain or harsh operational conditions in outdoor environmental monitoring, introduces further complexity.

The problem in designing the placement of a number of sensor nodes within the RoI has been a very attractive scientific exploration. A number of studies has been carried out in this research area within the last few decades. As an overview, Younis and Akkaya presented a comprehensive survey of strategies and techniques in sensor networks deployment prior to 2008 [158, 159]. The following sub-section presents some of the interesting studies which have been done related to the effort in designing ESN.

2.5.1 Challenges in ESN Design

Deploying a sophisticated equipment in am unattended environment has never been an easy task. There are numbers of factors required to be considered [156]. Following are some common challenges in the deployment of ESN:

2.5.1.1 Limitation in Resources

The design and implementation of ESN are limited by three main resources: energy, storage, and computation power [66, 129]. Most sensor nodes are relying on battery power to operate and their operational period and frequency of measurement are constrained to the capacity of the battery being used. As sensor nodes operate on limited battery power, energy usage is a very important concern in ESN. In order to overcome such limitation, some sensor nodes are equipped with solar panel to harvest energy from the sun and stored it into a rechargeable battery [7, 140]. The frequency and the period of monitoring are also limited by the storage capacity of the sensor node. Sensor nodes with low storage capacity may have to lower their measurement frequency or they can also transmit their data to their base station more frequent. The increase in data transmission frequency will also increase the energy consumption. In terms of computational capability, some sensor nodes are able to do a simple in-situ data pre-processing with a very limited computational power. Majority of the nodes are manufactured with sensing capability only and data processing will be handled by base station or sensor network server.

2.5.1.2 Deployment Area

In terms of deployment area, some ESN applications require to cover a very large region which often involve high amount of sensor nodes or few number of nodes with spread deployment locations. The increase in the number of nodes will also lead to the increase in not only deployment cost but also maintenance cost. Some deployment areas are remote and isolated which often difficult to reach. In this case, the deployment and maintenance of sensor nodes can only be done by engineer or technician with specific skill set required to reach the deployment areas. In addition, some applications require the network to be deployed in extreme and hostile regions such as glacier [111], active volcano [152], and battle field [97].

2.5.1.3 Harsh Environmental Condition

Harsh environmental condition is a common and inevitable challenge in most outdoor monitoring applications. Sensor nodes are prone to numerous extreme environmental events, namely heavy rain, extreme temperature variations, storm. In addition, the activity of unexpected visitors such as birds and other wild animals may also break the sensor nodes or cause failure in communication link [6]. Long exposure to such harsh environmental condition could degrade the performance and the overall lifetime of an ESN. In this case, regular visit and maintenance, which also include sensor calibration, are required in order to ensure the performance and the operational of the ESN [44].

2.5.2 Deployment Strategy

There are two schema in sensor node deployment strategy: random deployment and deterministic deployment [11]. The decision related to the deployment strategy is mainly depend on the application of the ESN and the environmental condition where the networks is going to be deployed [129, 165].

Some ESN applications require a random deployment of sensor nodes as the only viable option. This is commonly found in the ESN application with a harsh environmental condition such as region which just been hit by a disaster or in the battlefield [141]. In such hostile environment, deterministic deployment is considered high risk and infeasible. Dropping sensor nodes using a helicopter/aeroplane above the ground or using a grenade launcher is more viable. In random deployment, the opportunity to optimise the deployment objectives is very limited. This is due to the nature of the deployment where there is less to no control in the placement of each node. Random spreading of sensor nodes is expected, although the node density and the level of redundancy could be controlled to some degree in order to achieve its deployment objectives [16, 148, 149].

Deterministic placement of sensor nodes often be considered as necessary especially when dealing with expensive sensor nodes, where the performance of the network is significantly affected by the position of the nodes. Deterministic approach is also viable for an ESN deployment which aims for a relatively long period of observation [67, 95]. Such deployment strategy gives more opportunity for optimisation in order to meet desired deployment objectives.

2.5.3 Deployment Objectives

Sensor nodes placement significantly impacts the effectiveness of an ESN and the efficiency of its operation [17, 60, 113]. The design process in ESN deployment is motivated by the need to maximise certain deployment objectives using the least amount of sensor nodes. Forming an optimised sensor nodes placement is not an easy problem, without exception for deterministic deployment, and it has been proven to be NP-hard (Non-deterministic Polynomial-time hard) for most formulations of sensor deployment [159]. This section presents the deployment objectives which have been explored in a number of studies in relation to ESN design.

Maximising the spatial coverage is a well known ESN deployment objective and has been studied the most. Such deployment objective is motivated by the idea that a complete knowledge of the RoI could be obtained by covering every location within the RoI. The ratio of the covered area to the overall RoI is commonly utilised as the metric to quantify the quality of the coverage. A perfect disk sensing coverage is adopted in several studies, where a sensor node is located at the center with equal radius in its sensing range [76, 103, 106, 150]. The disk sensing range assumption may not be suitable for some real world applications which require a high accurate scenarios, therefore some studies started to explore an irregular polygon sensing coverage [18].

In the early studies in ESN design, network connectivity was not considered to be included as an objective in ESN deployment. This was due to the assumption that the transmission range of a sensor node is always higher compared to its sensing range [159]. Therefore, a good network connectivity was always expected in every deployment with sufficient sensing coverage. Unfortunately this is not always the case, there are some applications where the communication range is limited. In this case, some issue related to network connectivity may occur where a certain level of redundancy in coverage is unavoidable in order to maintain the connectivity. Therefore, efforts in maximising the network connectivity are commonly conducted in conjunction with maximising the sensing coverage [83, 124, 138].

In some applications, the placement of sensor nodes would significantly impact the lifetime of the ESN. This is mainly happen due to the variations in node density which eventually lead to unbalanced data communication load, where some part of the network would have higher traffic compared to the rest of the network. This condition would result to a faster energy dissipation in some of the nodes within the network. On the contrary, in some other applications, a uniform sensor node distribution could rise another network lifetime related issues. The nodes deployed close to the base station would suffer from a rapid energy depletion and thus shorten the network lifetime [154, 160]. Sensor nodes deployment with the objective of maximising the network lifetime has been considered as one of the very interesting subject in ESN deployment and there are numbers of studies have been conducted [9, 27, 94].

Data fidelity is another important objective in ESN deployment. The term data fidelity is used in this context in order to avoid an ambiguity with data quality. Data fidelity refers to the degree to which the measured and collected data could capture and reproduce the state and conditions of its environment. In other words, fidelity could also be defined as the degree of similarity between the modeled environmental data which is constructed by the ESN and the real environmental condition [65, 70]. Distortion in the environment model associated with ESN data is unavoidable. In this case, the model could be improved by deploying more sensor nodes within the RoI, where more sampled data will be collected. However, this approach is undesirable due to the increase in cost. Alternatively, the sensor nodes need to be deployed in a certain position which lead to a minimum distortion. Such effort is formulated as the problem in finding the optimal nodes position to meet a desired level of data fidelity [88, 89, 164].

Chapters 3 and 4 have been removed for copyright or proprietary reasons.

They have been published as: Budi, S., de Souza, P., Timms, G., Malhotra, V., Turner, P., 2015. Optimisation in the design of environmental sensor networks with robustness consideration, Sensors, 15(12), 29765–29781

Budi, S., de Souza, P., Timms, G., Susanto, F., Malhotra, V., Turner, P., 2018. Mobile platform sampling for designing environmental sensor networks. Environmental monitoring and assessment 190(3), 1-15

Budi, S., Susanto, F., de Souza, P., Timms, G., Malhotra, V., Turner, P., 2018. In search for a robust design of environmental sensor networks, Environmental technology, 3996), 683-693

Chapter 5

Discussions and Conclusions

Our environment influences many different aspects in our life through a range of parameters. A useful set of measurements have been developed to measure and to record the state of each of these parameters. These measurements, if properly conducted, would help us to better understand our environment. Air temperature measurement, as an example, permits a useful information related to temperature changes in a region over period of time and how the changes are correlated to other regions. A good understanding on the air temperature behaviour in a region would benefit to support decision making in the sectors influenced by the changes in air temperature, namely agriculture and forestry.

Although the work presented in this thesis focuses on air temperature as the parameter of interest, the proposed method could also be applied to other environmental parameters. There are numerous interesting measurement parameters in the environment that are necessary in order to have a meaningful and useful understanding of the environment. Apart from air temperature monitoring, sensing water-related parameters is another important task since it has direct influence on our life. Relative humidity, which determines the level of moisture in the air, has been observed along with air temperature. Some studies have revealed interesting correlations between air temperature and relative humidity to support weather forecasting. Rainfall measurements are useful records of the environment particularly in agriculture area since most crop productions are highly impacted by rainfall. Wind related measurements (e.g., wind speed and wind direction) along with atmospheric pressure measurements have also been extensively studied as as they can significantly contribute to weather forecasting. Solar radiation has also acquired attention in recent years especially in the region closer to south pole due to the building of ozone hole and its impact on health and environmental risks. Sensing and predicting of such environmental patterns is of great interest for countries like Australia.

This chapter presents and discusses several key components in this thesis which contributes to the body of knowledge in the area of environmental sensor network design. Some fundamental limitations of the study are also described in this chapter. Further, this chapter also discusses some possible research works that can be explored on top of the work presented in this thesis.

5.1 Research Contributions

Designing is a critical process prior to the deployment of an Environmental Sensor Network (ESN). Careful ESN design would result in a network fit for its purpose. In this thesis, there are two key parameters in the design of ESN: number of sensor nodes and the placement of the nodes in the Region of Interest (RoI).

The current studies in ESN design is mainly focusing on reducing costs, which is directly influenced by the number of sensor nodes being deployed. Redundant nodes within an ESN are considered as inefficient and should be minimised. Unfortunately, sensor failure is commonly found within the operational period of an ESN. The failure in sensor could significantly degrade the effectiveness of the ESN. The network may no longer be able to provide a service which fit for its purpose. Introducing more sensor nodes into an ESN would improve the robustness of the networks; however, it would also increase the redundancy of the network, which is considered as inefficient in most design practices. In addition, robustness aspect has not been sufficiently considered in the current practice of ESN design.

The main contribution of this work is in the development of an alternative method in ESN design which aims to fill the gap in current ESN design methods. In the proposed method, redundancy is not considered as primary factor to be optimised (i.e., minimised), instead redundancy is considered as a factor to be balanced with robustness. The proposal was achieved through five key research components, which are aligned with the research objectives of this study (as described in Section 1.3):

- **Measure of Representativeness**
  ESN is deployed in a particular region with the purpose to understand certain environmental phenomena within the region. It is expected that an ESN would generate data which best represents the region. In this study, the effort to quantify the representativeness of an ESN is conducted by calculating the error/difference in the average spatial temperature (over certain period of time) between the actual data and the one yieled by the ESN (as described in Section 3.5.1). In this case ESN with lower error has better representativeness compared to the one with higher error. Such measure of representativeness is considerably simple and fast to compute, however, it is poor in

terms of spatial resolution. In order to address the need for a higher spatial resolution, another measurement is proposed. The representativeness of an ESN is measured by calculating the error/difference in spatial temperature (over certain period of time) between the actual data and the interpolated spatial temperature data (as described in Section 3.5.2). The more representative ESN would yield a lower interpolation error compared to the one which is less representative.

- **ESN Design Optimisation**
  The process of finding the location for sensor nodes placement in this study is defined as an optimisation problem. The RoI is mapped as a two dimensional space and the placement of a given number of sensor nodes is formulated as a decision variable. In this type of deployment, the search space is considerably large with each position yielding different levels of representativeness. The growth in search space is significantly increased with a larger area of deployment. In addition, the increase in the number of sensor nodes will also increase the number of possible placements (as described in Section 3.2). Considering the size of the search space, Evolutionary Algorithm (EA) is employed in this study to find the optimum placement of sensor nodes where representativeness is adopted as a fitness/objective function to be optimised. The algorithm is able to handle a large search space and is also capable of avoiding local optima during the search process (as described in Section 3.4). In this work, each ESN design is optimised exclusively according to the number of sensor nodes being assigned. In this case, ESN design with a higher number of nodes is not optimised based on the one with lower number of nodes. The experimental part in this study also indicates an improvement in representativeness with the increase in the number of sensor nodes. The improvement is considerably significant within the ESN design with a lower number of nodes and gradually decreases in the ESN with a higher number of nodes, where adding more nodes no longer leads to any meaningful improvement (as described in Section 4.1.1.1). This information would benefit in estimating the required number of sensor nodes to be deployed in order to suit the purpose of the networks. Further, by taking budget constraints into account, this information would assist the decision makers to conduct cost and benefit analysis.

- **Data Quality Issue and Robustness Consideration**
  ESN data quality issue was simulated in this work in order to analyse its impact on the representativeness of an ESN design (as described in Section 3.5.1). Two common ESN data quality issues were covered: gap and noise. Gaps in ESN data mainly occur due to sensor or communication failure, which introduces some missing values

in the data. In the case of noise, the sensor still produces some data, however, the data does not accurately represent the actual condition. Artificial gap and noise were introduced into the previously discovered optimum ESN design and their impact on the representativeness was analysed. ESN design which incorporated fewer number or sensor nodes in general is prone to a more significant degradation in representativeness compared to the ESN with more nodes (as described in Section 4.1.1.2 and 4.1.1.4). This analysis would benefit the decision maker in determining the number of sensor nodes to be deployed and how far the representativeness of the ESN may suffer with the occurrence of gap/noise. In addition, robustness of an ESN was also considered in this work. Two techniques were employed to address the common ESN data quality issues: Spatial Regression Test (SRT) for gap filling and a simple temperature threshold as an automated data quality control (as described in Section 4.1.1.3 and 4.1.1.5). This robustness support is applied in conjunction with the data quality issue simulation in order to produce an overview of the impact of a certain level of quality issues to the ESN representativeness, as well as the effectiveness of the technique in promoting the robustness of the ESN.

- **Redundancy and Robustness Assessment**
  The effort in finding a balanced ESN design (in terms of redundancy and robustness) in this work is realised by formalising both redundancy and robustness as fitness/objective functions to be optimised (as described in Section 3.5.2). Redundancy is formulated as unnecessary deployment of sensor nodes in which their role can be handled by the neighbouring nodes using certain spatial data interpolation techniques. Whereas, robustness is formalised as how an ESN design could maintain the networks' performance while dealing with the loss or disruption of a node within the networks by utilising a temporal data interpolation technique. Since there are more than one criteria to be optimised, this approach leads to a multi-objective optimisation problem. Instead of having one single optimum ESN design, each number of sensor nodes would yield a different set of possible near-optimum placements of the nodes across the RoI. Each placement will produce a particular composition among two fitness/objective values (e.g., redundancy and robustness). Once a number of nodes has been decided, the next decision to be made is selecting an ESN design from a set of near-optimum designs (as described in Section 4.1.2.2 and 4.1.2.3). This method provides a number of ESN design options, allowing decision makers to have more control.

- **Mobile Data Sampling**
  Finding an ideal distribution of sensor nodes within a new region where historical

data is not available has never been a trivial task (as described in Section 3.6). In this case, an efficient data collection (i.e., data sampling) technique is needed. Data covering multiple years is desired and recommended to capture seasonal effects in a region. However, this would result in long delays and high costs of the deployment. In this study, a number of mobile platforms equipped with sensors were employed over periods of 30 days around the targeted dates to build a base knowledge of the RoI (as described in Section 3.6.2). Four dates that represent equinoxes and solstices provided convenient points for season identification. This enhances cost effectiveness as only a few mobile platforms need to be operated for data sampling. At the same time, the approach overcomes the errors that may result from extreme environmental events on the date of data collection. The sampled data is then enhanced with an interpolation technique in order to construct a complete sampling cube which aimed to substitute the absence of historical data (as described in Section 3.6.3).

5.2 Limitation of the Study

Apart from the contribution of this work to the body of knowledge, there are also number of limitations which bound this work. One of the major limitations was related to the dataset being used in this study. This study relied on SouthEsk Hydrodynamic model, targeting specifically on an hourly temperature data within one year period (2013). Other environmental properties such as wind speed/direction, relative humidity, solar radiation, or rainfall might have a unique and distinct characteristic compared to air temperature. The measure of representativeness in this study was also tailored for air temperature data. The measurement relied on the averaged value which would not be relevant in the case where the data is not normally distributed, such as with rain fall data. In this case, the proposed method could be ineffective dealing with environmental data which is not normally distributed.

Another notable limitation in this study is the assumption of a perfect flat surface where altitude/elevation of the RoI is neglected. This assumption would minimise the computational complexity since the spatial resolution is reduced. The low resolution in the space is then compensated with a higher temporal resolution where the air temperature is measured on an hourly basis. In this case, there is a possibility that certain environmental characteristics correlated to spatial distribution may not have been captured.

5.3 Direction of Future Research

There are a number of potential scientific questions that can be explored as future work based on the study presented in this thesis. This section is dedicated to highlight some of the potential future work. Air temperature has been chosen as the only environmental parameter in this study. However, this study can also be extended to other environmental parameters such as relative humidity and solar radiation. Theoretically, these two parameters are highly correlated with air temperature. In the situation where air temperature data is not captured, solar radiation and relative humidity data can potentially be exploited to estimate the air temperature value and fill the gap in the data. Extending this study to cover these two parameters may open a possibility for the development of new method to support robustness in the design of ESN.

Time Series Analysis (TSA) technique such as Auto-Regressive Integrated Moving Average (ARIMA) [90, 145] has the potential to be integrated into the methodology in this study. ARIMA is a well-established TSA technique which is mainly used in forecasting. The technique may also be applied to estimate missing values or gap in the data, which is commonly found in environmental monitoring practices. This capability introduces an opportunity for ARIMA to be applied as an extension for Spatial Regression Test (SRT) in this study (for gap filling application). Apart from its potential capability in estimating future values (including missing values), ARIMA requires a higher computational resource compared to SRT. In addition, ARIMA mainly relies on temporal dimension to estimate data (spatial dimension is not considered). However, the lack of spatial aspect in ARIMA can be tackled by the support of other interpolation technique like SRT, which utilises spatial dimension in its estimation.

Configuration for number of stationary sensor nodes has been the main focus of this thesis. Stationary ESNs are known for their capability in producing environmental monitoring data with high temporal resolution. Compared to mobile ESNs, stationary ESNs are relatively poor in terms of spatial resolution. Depending on the size of the region being monitored, a certain degree of improvement in spatial resolution could be achieved by introducing more sensor nodes into the network. Forming a hybrid ESN, where the number of mobile sensor nodes are operated in conjunction with the stationary sensor nodes, could be adopted as an alternative to have a balance in temporal and spatial resolution. Designing a hybrid ESN is another interesting area to be explored on the basis of the study presented in this thesis.

References

[1] J. Aalto, P. Pirinen, J. Heikkinen, and A. Venalainen. Spatial interpolation of monthly climate data for Finland: comparing the performance of kriging and generalized additive models. *Theoretical and Applied Climatology*, 112(1-2):99–111, Jul 2012. doi: 10.1007/s00704-012-0716-9.

[2] F. Adolphi, R. Muscheler, A. Svensson, A. Aldahan, G. Possnert, J. Beer, J. Sjolte, S. Björck, K. Matthes, and R. Thiéblemont. Persistent link between solar activity and Greenland climate during the Last Glacial Maximum. *Nature Geoscience*, 7(9): 662–666, Sept 2014. doi: 10.1038/ngeo2225.

[3] I. F. Akyildiz and M. C. Vuran. *Wireless Sensor Networks*. Wiley, Chichester, West Sussex, U.K. ; Hoboken, NJ, 1 edition edition, Aug 2010. ISBN 978-0-470-03601-3.

[4] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. A survey on sensor networks. *IEEE Communications Magazine*, 40(8):102–114, Aug 2002. doi: 10.1109/MCOM.2002.1024422.

[5] J. N. Al-Karaki and A. E. Kamal. Routing techniques in wireless sensor networks: a survey. *IEEE Wireless Communications*, 11(6):6–28, Dec 2004. doi: 10.1109/MWC.2004.1368893.

[6] F. M. Al-Turjman, H. S. Hassanein, and M. A. Ibnkahla. Efficient deployment of wireless sensor networks targeting environment monitoring applications. *Computer Communications*, 36(2):135–148, Jan 2013. doi: 10.1016/j.comcom.2012.08.021.

[7] C. Alippi and C. Galperti. An Adaptive System for Optimal Solar Energy Harvesting in Wireless Sensor Network Nodes. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 55(6):1742–1750, Jul 2008. doi: 10.1109/TCSI.2008.922023.

[8] T. Arampatzis, J. Lygeros, and S. Manesis. A Survey of Applications of Wireless Sensors and Wireless Sensor Networks. In *Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation Intelligent Control, 2005.*, pages 719–724, June 2005. doi: 10.1109/.2005.1467103.

[9] B. O. Ayinde and A. Y. Barnawi. Differential evolution based deployment of wireless sensor networks. In *2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA)*, pages 131–137, Nov 2014. doi: 10.1109/AICCSA.2014.7073189.

[10] R. Badia-Melis, L. Ruiz-Garcia, J. Garcia-Hierro, and J. I. R. Villalba. Refrigerated Fruit Storage Monitoring Combining Two Different Wireless Sensing Technologies: RFID and WSN. *Sensors*, 15(3):4781–4795, Feb 2015. doi: 10.3390/s150304781.

[11] P. Balister and S. Kumar. Random vs. Deterministic Deployment of Sensors in the Presence of Failures and Placement Errors. In *IEEE INFOCOM 2009*, pages 2896–2900, Apr 2009. doi: 10.1109/INFCOM.2009.5062254.

[12] B. J. Benson, B. J. Bond, M. P. Hamilton, R. K. Monson, and R. Han. Perspectives on next-generation technology for environmental sensor networks. *Frontiers in Ecology and the Environment*, 8(4):193–200, Jun 2009. doi: 10.1890/080130.

[13] V. Bhandary, A. Malik, and S. Kumar. Routing in Wireless Multimedia Sensor Networks: A Survey of Existing Protocols and Open Research Issues. *Journal of Engineering*, 2016, Apr 2016. doi: 10.1155/2016/9608757.

[14] G. Bitella, R. Rossi, R. Bochicchio, M. Perniola, and M. Amato. A novel low-cost open-hardware platform for monitoring soil water content and multiple soil-air-vegetation parameters. *Sensors*, 14(10):19639–19659, oct 2014. doi: 10.3390/s141019639.

[15] P. d. A. Borges, J. Franke, Y. M. T. Anunciacao, H. Weiss, and C. Bernhofer. Comparison of spatial interpolation methods for the estimation of precipitation distribution in Distrito Federal, Brazil. *Theoretical and Applied Climatology*, 123(1-2):335–348, Jan 2015. doi: 10.1007/s00704-014-1359-9.

[16] N. Boudriga. On a Controlled Random Deployment WSN-Based Monitoring System Allowing Fault Detection and Replacement. *International Journal of Distributed Sensor Networks*, 2014:e101496, Apr 2014. doi: 10.1155/2014/101496.

[17] A. Boukerche, editor. *Algorithms and protocols for wireless sensor networks*. Wiley series on parallel and distributed computing. Wiley, Hoboken, NJ, 2009. ISBN 978-0-471-79813-2.

[18] A. Boukerche, X. Fei, and R. B. Araujo. A Coverage-Preserving and Hole Tolerant Based Scheme for the Irregular Sensing Range in Wireless Sensor Networks. In *IEEE Globecom 2006*, pages 1–5, Nov 2006. doi: 10.1109/GLOCOM.2006.485.

[19] S. Budi, P. de Souza, G. Timms, V. Malhotra, and P. Turner. Optimisation in the Design of Environmental Sensor Networks with Robustness Consideration. *Sensors*, 15(12):29765–29781, Nov 2015. doi: 10.3390/s151229765.

[20] S. Budi, F. Susanto, P. de Souza, G. Timms, V. Malhotra, and P. Turner. In search for a robust design of environmental sensor networks. *Environmental Technology*, 39(6): 1–11, Mar 2017. doi: 10.1080/09593330.2017.1310303.

[21] S Budi, P de Souza, G Timms, F Susanto, V Malhotra, and P Turner. Mobile platform sampling for designing environmental sensor networks. *Environmental Monitoring and Assessment*, 190(3):130–144, Mar 2018. doi: 10.1007/s10661-018-6510-0.

[22] L. T. Bui. *Multi-Objective Optimization in Computational Intelligence: Theory and Practice*. IGI Global, Hershey, 1 edition edition, May 2008. ISBN 978-1-59904-498-9.

[23] D. K. Bums, P. N. Kember, S. Taylor, and E. W. Williams. Air pollution monitoring and the role of semiconductor sensors. *Environmental Technology Letters*, 1(5):259–272, May 1980. doi: 10.1080/09593338009383975.

[24] P. A. Burrough, R. A. McDonnell, and C. D. Lloyd. *Principles of Geographical Information Systems*. Oxford University Press, Oxford ; New York, 3 edition edition, Jun 2015. ISBN 978-0-19-874284-5.

[25] W. Cai, S. Borlace, M. Lengaigne, P. van Rensch, M. Collins, G. Vecchi, A. Timmermann, A. Santoso, M. J. McPhaden, L. Wu, M. H. England, G. Wang, E Guilyardi, and F. Jin. Increasing frequency of extreme El Nino events due to greenhouse warming. *Nature Climate Change*, 4(2):111–116, Feb 2014. doi: 10.1038/nclimate2100.

[26] J. L. Campbell, L. E. Rustad, J. H. Porter, J. R. Taylor, E. W. Dereszynski, J. B. Shanley, C. Gries, D. L. Henshaw, M. E. Martin, W. M. Sheldon, and E. R. Boose. Quantity is Nothing without Quality: Automated QA/QC for Streaming Environmental Sensor Data. *BioScience*, 63(7):574–585, Jul 2013. doi: 10.1525/bio.2013.63.7.10.

[27] M. Cardei and D. Z. Du. Improving Wireless Sensor Network Lifetime Through Power Aware Organization. *Wirel. Netw.*, 11(3):333–340, May 2005. doi: 10.1007/s11276-005-6615-6.

[28] C. J. Carmona, P. Gonzalez, M. J. del Jesus, and F. Herrera. Overview on evolutionary subgroup discovery: analysis of the suitability and potential of the search performed by evolutionary algorithms. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 4(2):87–103, Mar 2014. doi: 10.1002/widm.1118.

[29] F. Carse, M. J. Martin, A. Sellar, and E. W. Blockley. Impact of assimilating temperature and salinity measurements by animal-borne sensors on FOAM ocean model fields. *Quarterly Journal of the Royal Meteorological Society*, 141(693):2934–2943, Oct 2015. doi: 10.1002/qj.2613.

[30] D. Chen, T. Lian, C. Fu, M. A. Cane, Y. Tang, R. Murtugudde, X. Song, Q. Wu, and L. Zhou. Strong influence of westerly wind bursts on El Nino diversity. *Nature Geoscience*, 8(5):339–345, May 2015. doi: 10.1038/ngeo2399.

[31] P. Chen, Y. Li, Y. Su, X. Chen, and Y. Jiang. Review of AUV Underwater Terrain Matching Navigation. *The Journal of Navigation*, 68(6):1155–1172, Nov 2015. doi: 10.1017/S0373463315000429.

[32] C. Y. Chong and S. P. Kumar. Sensor networks: evolution, opportunities, and challenges. *Proceedings of the IEEE*, 91(8):1247–1256, Aug 2003. doi: 10.1109/JPROC.2003.814918.

[33] M. Claverie, J. L. Matthews, E. F. Vermote, and C. O. Justice. A 30+ year avhrr lai and fapar climate data record: Algorithm description and validation. *Remote Sensing*, 8(3):263, Mar 2016. doi: 10.3390/rs8030263.

[34] C. C. Coello. Evolutionary multi-objective optimization: a historical view of the field. *IEEE Computational Intelligence Magazine*, 1(1):28–36, Feb 2006. doi: 10.1109/MCI.2006.1597059.

[35] C. C. Coello, G. B. Lamont, and D. A. van Veldhuizen. *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer, New York, 2nd edition edition, Sept 2007. doi: 10.1007/978-0-387-36797-2.

[36] M. Comboul and R. Ghanem. Value of Information in the Design of Resilient Water Distribution Sensor Networks. *Journal of Water Resources Planning and Management*, 139(4):449–455, 2013. doi: 10.1061/(ASCE)WR.1943-5452.0000259.

[37] P. Corke, T. Wark, R. Jurdak, W. Hu, P. Valencia, and D. Moore. Environmental Wireless Sensor Networks. *Proceedings of the IEEE*, 98(11):1903–1917, Nov 2010. doi: 10.1109/JPROC.2010.2068530.

[38] N. Cressie. Spatial prediction and ordinary kriging. *Mathematical Geology*, 20(4): 405–421, May 1988. doi: 10.1007/BF00892986.

[39] T. E. Cropper and P. E. Cropper. A 133-Year Record of Climate Change and Variability from Sheffield, England. *Climate*, 4(3):46, Sep 2016. doi: 10.3390/cli4030046.

[40] K. A. De Jong. Parameter Setting in EAs: a 30 Year Perspective. In F. G. Lobo, C. F. Lima, and Z. Michalewicz, editors, *Parameter Setting in Evolutionary Algorithms*, number 54 in Studies in Computational Intelligence, pages 1–18. Springer Berlin Heidelberg, 2007. doi: 10.1007/978-3-540-69432-8_1.

[41] K. A. De Jong and W. M. Spears. An analysis of the interacting roles of population size and crossover in genetic algorithms. In H. P. Schwefel and R. Manner, editors, *Parallel Problem Solving from Nature*, number 496 in Lecture Notes in Computer Science, pages 38–47. Springer Berlin Heidelberg, Oct 1990. doi: 10.1007/BFb0029729.

[42] K. Deb. *Multi-Objective Optimization Using Evolutionary Algorithms*. Wiley, 2 edition edition, Oct 2008. ISBN 978-0-470-74361-4.

[43] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2): 182–197, Apr 2002. doi: 10.1109/4235.996017.

[44] C. D'Este, P. de Souza, C. Sharman, and S. Allen. Relocatable, automated cost-benefit analysis for marine sensor network design. *Sensors*, 12(3):2874–2898, Mar 2012. doi: 10.3390/s120302874.

[45] A. Di Piazza, F. L. Conti, L. V. Noto, F. Viola, and G. La Loggia. Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for sicily, italy. *International Journal of Applied Earth Observation and Geoinformation*, 13(3):396–408, Jun 2011. doi: 10.1016/j.jag.2011.01.005.

[46] H. J. Diamond, T. R. Karl, M. A. Palecki, C. B. Baker, J. E. Bell, R. D. Leeper, D. R. Easterling, J. H. Lawrimore, T. P. Meyers, M. R. Helfert, G. Goodge, and P. W. Thorne. U.S. Climate Reference Network after One Decade of Operations: Status and Assessment. *Bulletin of the American Meteorological Society*, 94(4):485–498, Apr 2013. doi: 10.1175/BAMS-D-12-00170.1.

[47] J. Dong, G. Wang, H. Yan, J. Xu, and X. Zhang. A survey of smart water quality monitoring system. *Environmental Science and Pollution Research*, 22(7):4893–4906, Jan 2015. doi: 10.1007/s11356-014-4026-x.

[48] A. Donohoe and D. S. Battisti. The Seasonal Cycle of Atmospheric Heating and Temperature. *Journal of Climate*, 26(14):4962–4980, Jan 2013. doi: 10.1175/JCLI-D-12-00713.1.

[49] I. Durre, M. J. Menne, B. E. Gleason, T. G. Houston, and R. S. Vose. Comprehensive Automated Quality Assurance of Daily Surface Observations. *Journal of Applied Meteorology and Climatology*, 49(8):1615–1633, Feb 2010. doi: 10.1175/2010JAMC2375.1.

[50] R. Dutta, A. Das, and J. Aryal. Big data integration shows Australian bush-fire frequency is increasing significantly. *Royal Society open science*, 3(2):150241, Feb 2016. doi: 10.1098/rsos.150241.

[51] B. Elen, J. Peters, M. van Poppel, N. Bleux, J. Theunis, M. Reggente, and A. Standaert. The aeroflex: A bicycle for mobile air quality measurements. *Sensors*, 13(1):221–240, Dec 2012. doi: 10.3390/s130100221.

[52] M. Emadi and M. Baghernejad. Comparison of spatial interpolation techniques for mapping soil pH and salinity in agricultural coastal areas, northern Iran. *Archives of Agronomy and Soil Science*, 60(9):1315–1327, Sept 2014. doi: 10.1080/03650340.2014.880837.

[53] G. Fan, R. Wang, H. Huang, L. Sun, and C. Sha. Coverage-guaranteed sensor node deployment strategies for wireless sensor networks. *Sensors*, 10(3):2064–2087, Mar 2010. doi: 10.3390/s100302064.

[54] S. Faye and C. Chaudet. Characterizing the Topology of an Urban Wireless Sensor Network for Road Traffic Management. *IEEE Transactions on Vehicular Technology*, 65(7):5720–5725, Jul 2016. doi: 10.1109/TVT.2015.2465811.

[55] E. Felemban, F. K. Shaikh, U. M. Qureshi, A. A. Sheikh, and S. B. Qaisar. Underwater Sensor Network Applications: A Comprehensive Survey. *International Journal of Distributed Sensor Networks*, 11(11):896832, Nov 2015. doi: 10.1155/2015/896832.

[56] F. J. Ferrández-Pastor, J. M. García-Chamizo, M. Nieto-Hidalgo, J. Mora-Pascual, and J. Mora-Martínez. Developing Ubiquitous Sensor Network Platform Using Internet of Things: Application in Precision Agriculture. *Sensors*, 16(7):1141, Jul 2016. doi: 10.3390/s16071141.

[57] D. B. Fogel. *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*. Wiley-IEEE Press, Hoboken, N.J, 3 edition edition, Dec 2005. ISBN 978-0-471-66951-7.

[58] F. A. Fortin, F. M. D. Rainville, M. D. Gardner, M. Parizeau, and C. Gagné. Deap: Evolutionary algorithms made easy. *Journal of Machine Learning Research*, 13:2171–2175, Jul 2012. ISSN 1533-7928.

[59] P. Foukal, C. Fröhlich, H. Spruit, and T. M. L. Wigley. Variations in solar luminosity and their effect on the Earth's climate. *Nature*, 443(7108):161–166, Sept 2006. doi: 10.1038/nature05072.

[60] F. Franceschini, editor. *Distributed large-scale dimensional metrology: new insights*. Springer, London, 2011. ISBN 978-0-85729-542-2.

[61] G. Furquim, G. Pessin, B. S. Faiçal, E. M. Mendiondo, and J. Ueyama. Improving the accuracy of a flood forecasting model by means of machine learning and chaos theory. *Neural Computing and Applications*, 27(5):1129–1141, Jul 2016. doi: 10.1007/s00521-015-1930-z.

[62] A. R. Ganguly and K. Steinhaeuser. Data mining for climate change and impacts. In *IEEE International Conference on Data Mining Workshops, 2008. ICDMW '08*, pages 385–394, Dec 2008. doi: 10.1109/ICDMW.2008.30.

[63] D. Gavalas and M. Kenteris. A web-based pervasive recommendation system for mobile tourist guides. *Personal and Ubiquitous Computing*, 15(7):759–770, Oct 2011. doi: 10.1007/s00779-011-0389-x.

[64] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Professional, Reading, Mass, 1 edition edition, Jan 1989. ISBN 978-0-201-15767-3.

[65] D. C. Gross. Report from the fidelity implementation study group. In *Fall Simulation Interoperability Workshop Papers*, 1999.

[66] V. C. Gungor and G. P. Hancke. Industrial wireless sensor networks: Challenges, Design Principles, and Technical Approaches. *IEEE Transactions on Industrial Electronics*, 56(10):4258–4265, Oct 2009. doi: 10.1109/TIE.2009.2015754.

[67] V. C. Gungor, B. Lu, and G. P. Hancke. Opportunities and challenges of wireless sensor networks in smart grid. *IEEE Transactions on Industrial Electronics*, 57(10): 3557–3564, Oct 2010. doi: 10.1109/TIE.2009.2039455.

[68] J. K. Hart and K. Martinez. Environmental sensor networks: A revolution in the earth system science? *Earth-Science Reviews*, 78(3–4):177–191, Oct 2006. doi: 10.1016/j.earscirev.2006.05.001.

[69] M. Hautefeuille, B. O'Flynn, F. H. Peters, and C. O'Mahony. Development of a Microelectromechanical System (MEMS)-Based Multisensor Platform for Environmental Monitoring. *Micromachines*, 2(4):410–430, Nov 2011. doi: 10.3390/mi2040410.

[70] R. T. Hays and M. J. Singer. *Simulation fidelity in training system design: bridging the gap between reality and training*. Recent research in psychology. Springer, New York, 1989. ISBN 978-0-387-96846-9 978-3-540-96846-7.

[71] E. Heinz, P. Kraft, C. Buchen, H. G. Frede, E. Aquino, and L. Breuer. Set up of an automatic water quality sampling system in irrigation agriculture. *Sensors*, 14(1): 212–228, Dec 2013. doi: 10.3390/s140100212.

[72] J. S. Horsburgh, A. S. Jones, D. K. Stevens, D. G. Tarboton, and N. O. Mesner. A sensor network for high frequency estimation of water quality constituent fluxes using surrogates. *Environmental Modelling & Software*, 25(9):1031–1044, Sept 2010. doi: 10.1016/j.envsoft.2009.10.012.

[73] T. D. Houston and R. Hiederer. Applying quality assurance procedures to environmental monitoring data: a case study. *Journal of Environmental Monitoring*, 11(4): 774–781, Apr 2009. doi: 10.1039/B818274B.

[74] B. Hu, W. Chen, and J. Zhou. High performance flexible sensor based on inorganic nanomaterials. *Sensors and Actuators B: Chemical*, 176:522–533, Jan 2013. doi: 10.1016/j.snb.2012.09.036.

[75] B. Huang, P. Fery, L. Xue, and Y. Wang. Seeking the Pareto front for multiobjective spatial optimization problems. *International Journal of Geographical Information Science*, 22(5):507–526, May 2008. doi: 10.1080/13658810701492365.

[76] C. F. Huang and Y. C. Tseng. The coverage problem in a wireless sensor network. *Mob. Netw. Appl.*, 10(4):519–528, Aug 2005. doi: 10.1007/s11036-005-1564-y.

[77] K. G. Hubbard and J. You. Sensitivity Analysis of Quality Assurance Using the Spatial Regression Approach - A Case Study of the Maximum/Minimum Air Temperature. *Journal of Atmospheric and Oceanic Technology*, 22(10):1520–1530, Oct 2005. doi: 10.1175/JTECH1790.1.

[78] D. J. Isaak, S. Wollrab, D. Horan, and G. Chandler. Climate change effects on stream and river temperatures across the northwest u.s. from 1980–2009 and implications for salmonid fishes. *Climatic Change*, 113(2):499–524, Nov 2011. doi: 10.1007/s10584-011-0326-z.

[79] H. Ishibuchi, N. Akedo, and Y. Nojima. Behavior of Multiobjective Evolutionary Algorithms on Many-Objective Knapsack Problems. *IEEE Transactions on Evolutionary Computation*, 19(2):264–283, Apr 2015. doi: 10.1109/TEVC.2014.2315442.

[80] R. Jain. *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*. Wiley, 1991. ISBN 978-0-471-50336-1.

[81] F. Ji, Z. Wu, J. Huang, and E. P. Chassignet. Evolution of land surface air temperature trend. *Nature Climate Change*, 4(6):462–466, Jun 2014. doi: 10.1038/nclimate2223.

[82] J. Joo, J. Yim, and C. K. Lee. Protecting cultural heritage tourism sites with the ubiquitous sensor network. *Journal of Sustainable Tourism*, 17(3):397–406, May 2009. doi: 10.1080/09669580802582498.

[83] K Kar and S Banerjee. Node Placement for Connected Coverage in Sensor Networks. WiOpt'03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, March 2003. URL https://hal.inria.fr/inria-00466114.

[84] J Katzfey and M Thatcher. Ensemble one-kilometre forecasts for the south esk hydrological sensor web. In *Proceedings of the 19th International Congress on Modelling and Simulation, Perth, Australia*, pages 12–16, 2011. ISBN 978-0-9872143-1-7.

[85] R. Kays, M. C. Crofoot, W. Jetz, and M. Wikelski. Terrestrial animal tracking as an eye on life and planet. *Science*, 348(6240):aaa2478, Jun 2015. doi: 10.1126/science.aaa2478.

[86] K. Kim and H. Myung. Sensor node for remote monitoring of waterborne disease-causing bacteria. *Sensors*, 15(5):10569–10579, May 2015. doi: 10.3390/s150510569.

[87] N. Kotamäki, S. Thessler, J. Koskiaho, A. O. Hannukkala, H. Huitu, T. Huttula, J. Havento, and M. Järvenpää. Wireless in-situ Sensor Network for Agriculture and Water Monitoring on a River Basin Scale in Southern Finland: Evaluation from a Data User's Perspective. *Sensors*, 9(4):2862–2883, Apr 2009. doi: 10.3390/s90402862.

[88] A. Krause and C. Guestrin. Submodularity and Its Applications in Optimized Information Gathering. *ACM Trans. Intell. Syst. Technol.*, 2(4):32:1–32:20, Jul 2011. doi: 10.1145/1989734.1989736.

[89] A. Krause, C. Guestrin, A. Gupta, and J. Kleinberg. Near-optimal Sensor Placements: Maximizing Information While Minimizing Communication Cost. In *Proceedings of the 5th International Conference on Information Processing in Sensor Networks*, IPSN '06, pages 2–10, New York, NY, USA, 2006. ACM. doi: 10.1145/1127777.1127782.

[90] U. Kumar and V. K. Jain. ARIMA forecasting of ambient air pollutants (O3, NO, NO2 and CO). *Stochastic Environmental Research and Risk Assessment*, 24(5):751–760, July 2010. doi: 10.1007/s00477-009-0361-8.

[91] H. H. Lamb. *Climate: Present, Past and Future: Volume 1: Fundamentals and Climate Now*. Routledge, Abingdon, Oxon; New York, Nov 2012. ISBN 978-0-415-68222-0.

[92] S. Lan, M. Qilong, and J. Du. Architecture of wireless sensor networks for environmental monitoring. In *International Workshop on Education Technology and Training, 2008. and 2008 International Workshop on Geoscience and Remote Sensing. ETT and GRS 2008*, volume 1, pages 579–582, Dec 2008. doi: 10.1109/ETTandGRS.2008.219.

[93] K. Langendoen, A. Baggio, and O. Visser. Murphy loves potatoes: experiences from a pilot sensor network deployment in precision agriculture. In *Proceedings 20th IEEE International Parallel Distributed Processing Symposium*, pages 8 pp.–, Apr 2006. doi: 10.1109/IPDPS.2006.1639412.

[94] J. M. Lanza-Gutierrez and J. A. Gomez-Pulido. Assuming multiobjective metaheuristics to solve a three-objective optimisation problem for Relay Node deployment in Wireless Sensor Networks. *Applied Soft Computing*, 30:675–687, May 2015. doi: 10.1016/j.asoc.2015.01.051.

[95] M. A. Lazzara, G. A. Weidner, L. M. Keller, J. E. Thom, and J. J. Cassano. Antarctic Automatic Weather Station Program: 30 Years of Polar Observation. *Bulletin of the American Meteorological Society*, 93(10):1519–1537, Mar 2012. doi: 10.1175/BAMS-D-11-00015.1.

[96] H.C. Lee and K.H. Ke. Monitoring of Large-Area IoT Sensors Using a LoRa Wireless Mesh Network System: Design and Evaluation. *IEEE Transactions on Instrumentation and Measurement*, pages 1–11, 2018. doi: 10.1109/TIM.2018.2814082.

[97] S. H. Lee, S. Lee, H. Song, and H. S. Lee. Wireless sensor network design for tactical military applications : Remote large-scale environments. In *MILCOM 2009 - 2009 IEEE Military Communications Conference*, pages 1–7, Oct 2009. doi: 10.1109/MILCOM.2009.5379900.

[98] D. Li, J. J. Cao, and Y. Yao. Big data in smart cities. *Science China Information Sciences*, 58(10):1–12, Oct 2015. doi: 10.1007/s11432-015-5396-5.

[99] J. Li and A. D. Heap. *A review of spatial interpolation methods for environmental scientists*. Geoscience Australia Canberra, Canberra, 2008. URL http://www.ga.gov.au/metadata-gateway/metadata/record/gcat_68229.

[100] J. Li and A. D. Heap. A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. *Ecological Informatics*, 6 (3–4):228–241, Jul 2011. doi: 10.1016/j.ecoinf.2010.12.003.

[101] J. Li and A. D. Heap. Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling & Software*, 53:173–189, Mar 2014. doi: 10.1016/j.envsoft.2013.12.008.

[102] W. Li, S. Wu, M. Song, and X. Zhou. A scalable cyberinfrastructure solution to support big data management and multivariate visualization of time-series sensor observation data. *Earth Science Informatics*, 9(4):449–464, Nov 2016. doi: 10.1007/s12145-016-0267-1.

[103] X. Li, H. Frey, N. Santoro, and I. Stojmenovic. Strictly Localized Sensor Self-Deployment for Optimal Focused Coverage. *IEEE Transactions on Mobile Computing*, 10(11):1520–1533, Nov 2011. doi: 10.1109/TMC.2010.261.

[104] Q. Liao, L. Shi, Y. He, R. Li, Z. Su, A. Striegel, and Y. Liu. Visualizing Anomalies in Sensor Networks. In *Proceedings of the ACM SIGCOMM 2011 Conference*, SIGCOMM '11, pages 460–461, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0797-0. doi: 10.1145/2018436.2018521.

[105] Y. Liao, M. Mollineaux, R. Hsu, R. Bartlett, A. Singla, A. Raja, R. Bajwa, and R. Rajagopal. SnowFort: An Open Source Wireless Sensor Network for Data Analytics in Infrastructure and Environmental Monitoring. *IEEE Sensors Journal*, 14(12):4253–4263, Dec 2014. doi: 10.1109/JSEN.2014.2358253.

[106] F. Y. S. Lin and P. L. Chiu. A near-optimal sensor placement algorithm to achieve complete coverage-discrimination in sensor networks. *IEEE Communications Letters*, 9(1):43–45, Jan 2005. doi: 10.1109/LCOMM.2005.01027.

[107] N. Liu, W. Cao, Y. Zhu, J. Zhang, F. Pang, and J. Ni. The node deployment of intelligent sensor networks based on the spatial difference of farmland soil. *Sensors*, 15(11):28314–28339, Nov 2015. doi: 10.3390/s151128314.

[108] J. Lloret, M. Garcia, D. Bri, and S. Sendra. A wireless sensor network deployment for rural and forest fire detection and verification. *Sensors*, 9(11):8722–8747, Oct 2009. doi: 10.3390/s91108722.

[109] A. Malaver, N. Motta, P. Corke, and F. Gonzalez. Development and integration of a solar powered unmanned aerial vehicle and a wireless sensor network to monitor greenhouse gases. *Sensors*, 15(2):4072–4096, Feb 2015. doi: 10.3390/s150204072.

[110] K. Martinez, J. K. Hart, and R. Ong. Environmental sensor networks. *Computer*, 37 (8):50–56, Aug 2004. doi: 10.1109/MC.2004.91.

[111] K. Martinez, P. Padhy, A. Elsaify, G. Zou, A. Riddoch, J. K. Hart, and H. L. R. Ong. Deploying a sensor network in an extreme environment. In *IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC'06)*, volume 1, pages 8 pp.–, Jun 2006. doi: 10.1109/SUTC.2006.1636175.

[112] J. R. Martinez-de Dios, L. Merino, F. Caballero, and A. Ollero. Automatic forest-fire measuring using ground stations and unmanned aerial systems. *Sensors*, 11(6): 6328–6353, Jun 2011. doi: 10.3390/s110606328.

[113] M. J. McGrath and C. N. Scanaill. *Sensor Technologies: Healthcare, Wellness and Environmental Applications*. Apress, Berkeley, CA? : New York, 1 edition edition, Dec 2013. ISBN 978-1-4302-6013-4.

[114] S. Mejjaouli and R. F. Babiceanu. RFID-wireless sensor networks integration: Decision models and optimization of logistics systems operations. *Journal of Manufacturing Systems*, 35:234–245, Apr 2015. doi: 10.1016/j.jmsy.2015.02.005.

[115] I. A. Nalder and R. W. Wein. Spatial interpolation of climatic normals: test of a new method in the canadian boreal forest. *Agricultural and Forest Meteorology*, 92(4): 211–225, Dec 1998. doi: 10.1016/S0168-1923(98)00102-6.

[116] P. Ngatchou, A. Zarei, and A. El-Sharkawi. Pareto Multi Objective Optimization. In *Proceedings of the 13th International Conference on, Intelligent Systems Application to Power Systems*, pages 84–91, Nov 2005. doi: 10.1109/ISAP.2005.1599245.

[117] E. Onur, C. Ersoy, and H. Delic. How many sensors for an acceptable breach detection probability? *Computer Communications*, 29(2):173–182, Jan 2006. doi: 10.1016/j.comcom.2005.05.021.

[118] G. Pajares, A. Peruzzi, and P. Gonzalez-de Santos. Sensors in Agriculture and Forestry. *Sensors*, 13(9):12132–12139, Sept 2013. doi: 10.3390/s130912132.

[119] R. A. Peppler, C. N. Long, D. L. Sisterson, D. D. Turner, C. P. Bahrmann, S. W. Christensen, K. J. Doty, R. C. Eagan, T. D. Halter, M. D. Ivey, N. N. Keck, K. E. Kehoe, J. C. Liljegren, M. C. Macduff, J. H. Mather, R. A. McCord, J. W. Monroe, S. T. Moore, K. L. Nitschke, B. W. Orr, R. C. Perez, B. D. Perkins, S. J. Richardson, K. L. Sonntag, J. W. Voyles, and R. Wagener. An overview of arm program climate research facility data quality assurance. *The Open Atmospheric Science Journal*, 2(1): 192–216, Oct 2008. doi: 10.2174/1874282300802010192.

[120] F. J. Pierce and T. V. Elliott. Regional and on-farm wireless sensor networks for agricultural systems in eastern washington. *Computers and Electronics in Agriculture*, 61(1):32–43, Apr 2008. doi: 10.1016/j.compag.2007.05.007.

[121] J. H. Porter, E. Nagy, T. K. Kratz, P. Hanson, S. L. Collins, and P. Arzberger. New eyes on the world: Advanced sensors for ecology. *BioScience*, 59(5):385–397, May 2009. doi: 10.1525/bio.2009.59.5.6.

[122] J. H. Porter, P. C. Hanson, and C. C. Lin. Staying afloat in the sensor data deluge. *Trends in Ecology & Evolution*, 27(2):121–129, Feb 2012. doi: 10.1016/j.tree.2011.11.009.

[123] M. A. Rassam, M. A. Maarof, and A. Zainal. Adaptive and online data anomaly detection for wireless sensor systems. *Knowledge-Based Systems*, 60:44–57, Apr 2014. doi: 10.1016/j.knosys.2014.01.003.

[124] M. Rebai, M. Le berre, H. Snoussi, F. Hnaien, and L. Khoukhi. Sensor deployment optimization methods to achieve both coverage and connectivity in wireless sensor networks. *Computers & Operations Research*, 59:11–21, Jul 2015. doi: 10.1016/j.cor.2014.11.002.

[125] C. Reed, K. Buehler, and L. McKee. OGC Consensus: How Successful Standards Are Made. *ISPRS International Journal of Geo-Information*, 4(3):1693–1706, Sept 2015. doi: 10.3390/ijgi4031693.

[126] R. Rew and G. Davis. Netcdf: an interface for scientific data access. *IEEE Computer Graphics and Applications*, 10(4):76–82, Jul 1990. doi: 10.1109/38.56302.

[127] P. Ridao, M. Carreras, D. Ribas, P. J. Sanz, and G. Oliver. Intervention AUVs: The next challenge. *Annual Reviews in Control*, 40:227–241, 2015. doi: 10.1016/j.arcontrol.2015.09.015.

[128] S. Ripperger, D. Josic, M. Hierold, A. Koelpin, R. Weigel, M. Hartmann, R. Page, and F. Mayer. Automated proximity sensing in small vertebrates: design of miniaturized sensor nodes and first field tests in bats. *Ecology and Evolution*, 6(7):2179–2189, Apr 2016. doi: 10.1002/ece3.2040.

[129] K. Romer and F. Mattern. The design space of wireless sensor networks. *IEEE Wireless Communications*, 11(6):54–61, Dec 2004. doi: 10.1109/MWC.2004.1368897.

[130] E. Salamí, C. Barrado, and E. Pastor. UAV Flight Experiments Applied to the Remote Sensing of Vegetated Areas. *Remote Sensing*, 6(11):11051–11081, Nov 2014. doi: 10.3390/rs61111051.

[131] M. Shahbazi, J. Theau, and P. Ménard. Recent applications of unmanned aerial imagery in natural resource management. *GIScience & Remote Sensing*, 51(4):339–365, Jul 2014. doi: 10.1080/15481603.2014.926650.

[132] D. Shepard. A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 1968 23rd ACM National Conference*, ACM '68, pages 517–524, New York, NY, USA, 1968. ACM. doi: 10.1145/800186.810616.

[133] W. Shi, M. S. Wong, J. Wang, and Y. Zhao. Analysis of airborne particulate matter (pm2.5) over hong kong using remote sensing and gis. *Sensors*, 12(6):6825–6836, May 2012. doi: 10.3390/s120606825.

[134] C. W. Shih and C. H. Wang. Integrating wireless sensor networks with statistical quality control to develop a cold chain system in food industries. *Computer Standards & Interfaces*, 45:62–78, Mar 2016. doi: 10.1016/j.csi.2015.12.004.

[135] L. M. C. Silva, A. M. Salgado, and M. A. Z. Coelho. Development of an amperometric biosensor for phenol detection. *Environmental Technology*, 32(5):493–497, Apr 2011. doi: 10.1080/09593330.2010.504234.

[136] K. Sohraby, D. Minoli, and T. Znati. *Wireless Sensor Networks: Technology, Protocols, and Applications*. Wiley-Interscience, Hoboken, N.J, 1 edition edition, Apr 2007. ISBN 978-0-471-74300-2.

[137] M. V. Storey, B. van der Gaag, and B. P. Burns. Advances in on-line drinking water quality monitoring and early warning systems. *Water Research*, 45(2):741–747, Jan 2011. doi: 10.1016/j.watres.2010.08.049.

[138] Y. Sun, Z. Yu, J. Ge, B. Lin, and Z. Yun. On Deploying Wireless Sensors to Achieve Both Coverage and Connectivity. In *5th International Conference on Wireless Communications, Networking and Mobile Computing, 2009. WiCom '09*, pages 1–4, Sept 2009. doi: 10.1109/WICOM.2009.5300925.

[139] F. Susanto, S. Budi, P. de Souza, U. Engelke, and J. He. Design of Environmental Sensor Networks Using Evolutionary Algorithms. *IEEE Geoscience and Remote Sensing Letters*, 13(4):575–579, 2016. doi: 10.1109/LGRS.2016.2525980.

[140] J. Taneja, J. Jeong, and D. Culler. Design, Modeling, and Capacity Planning for Micro-solar Power Sensor Networks. In *Proceedings of the 7th International Conference on Information Processing in Sensor Networks*, IPSN '08, pages 407–418, Washington, DC, USA, 2008. IEEE Computer Society. doi: 10.1109/IPSN.2008.67.

[141] A. S. Tanenbaum, C. Gamage, and B. Crispo. Taking Sensor Networks from the Lab to the Jungle. *Computer*, 39(8):98–100, Aug 2006. doi: 10.1109/MC.2006.280.

[142] G. P. Timms, P. A. de Souza, L. Reznik, and D. V. Smith. Automated Data Quality Assessment of Marine. *Sensors*, 11(10):9589–9602, Oct 2011. doi: 10.3390/s111009589.

[143] K. E. Trenberth, J. T. Fasullo, G. Branstator, and A. S. Phillips. Seasonal aspects of the recent pause in surface warming. *Nature Climate Change*, 4(10):911–916, Oct 2014. doi: 10.1038/nclimate2341.

[144] G. Tsiourlis, S. Andreadakis, and P. Konstantinidis. Sithon: A wireless network of in situ optical cameras applied to the early detection-notification-monitoring of forest fires. *Sensors*, 9(6):4465–4482, Jun 2009. doi: 10.3390/s90604465.

[145] O. Valenzuela, I. Rojas, F. Rojas, H. Pomares, L. J. Herrera, A. Guillen, L. Marquez, and M. Pasadas. Hybridization of intelligent techniques and ARIMA models for time series prediction. *Fuzzy Sets and Systems*, 159(7):821–845, Apr 2008. doi: 10.1016/j.fss.2007.11.003.

[146] J. A. Vera-Gómez, A. Quesada-Arencibia, C. R. García, R. Suarez Moreno, and F. Guerra Hernandez. An Intelligent Parking Management System for Urban Areas. *Sensors*, 16(6):931, Jun 2016. doi: 10.3390/s16060931.

[147] C. von Lücken, B. Baran, and C. Brizuela. A survey on multi-objective evolutionary algorithms for many-objective problems. *Computational Optimization and Applications*, 58(3):707–756, Feb 2014. doi: 10.1007/s10589-014-9644-1.

[148] P. J. Wan and C. W. Yi. Coverage by randomly deployed wireless sensor networks. *IEEE Transactions on Information Theory*, 52(6):2658–2669, Jun 2006. doi: 10.1109/TIT.2005.862092.

[149] B. Wang, J. Zhu, L. T. Yang, and Y. Mo. Sensor Density for Confident Information Coverage in Randomly Deployed Sensor Networks. *IEEE Transactions on Wireless Communications*, 15(5):3238–3250, May 2016. doi: 10.1109/TWC.2016.2518689.

[150] X. Wang, G. Xing, Y. Zhang, C. Lu, R. Pless, and C. Gill. Integrated Coverage and Connectivity Configuration in Wireless Sensor Networks. In *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems*, SenSys '03, pages 28–39, New York, NY, USA, 2003. ACM. doi: 10.1145/958491.958496.

[151] R. Webster and M. A. Oliver. *Geostatistics for Environmental Scientists*. Wiley, 2nd edition, Nov 2007. ISBN 978-0-470-02858-2.

[152] G. Werner-Allen, K. Lorincz, M. Ruiz, O. Marcillo, J. Johnson, J. Lees, and M. Welsh. Deploying a wireless sensor network on an active volcano. *IEEE Internet Computing*, 10(2):18–25, Mar 2006. doi: 10.1109/MIC.2006.26.

[153] Y. Wu and M. Cardei. Multi-channel and cognitive radio approaches for wireless sensor networks. *Computer Communications*, 94:30–45, Nov 2016. doi: 10.1016/j.comcom.2016.08.010.

[154] K. Xu, H. Hassanein, G. Takahara, and Q. Wang. Relay node deployment strategies in heterogeneous wireless sensor networks: single-hop communication case. In *GLOBECOM '05. IEEE Global Telecommunications Conference, 2005.*, volume 1, pages 5 pp.–, Nov 2005. doi: 10.1109/GLOCOM.2005.1577051.

[155] X. Yang, T. Wang, J. Liang, G. Yao, and M. Liu. Survey on the novel hybrid aquatic–aerial amphibious aircraft: Aquatic unmanned aerial vehicle (AquaUAV). *Progress in Aerospace Sciences*, 74:131–151, Apr 2015. doi: 10.1016/j.paerosci.2014.12.005.

[156] J. Yick, B. Mukherjee, and D. Ghosal. Wireless sensor network survey. *Computer Networks*, 52(12):2292–2330, Aug 2008. doi: 10.1016/j.comnet.2008.04.002.

[157] J You, K. G. Hubbard, and S. Goddard. Comparison of methods for spatially estimating station temperatures in a quality control system. *International Journal of Climatology*, 28(6):777–787, May 2008. doi: 10.1002/joc.1571.

[158] M. Younis and K. Akkaya. Node positioning for increased dependability of Wireless Sensor Networks. In *Algorithms and Protocols for Wireless Sensor Networks*, pages 225–266. John Wiley & Sons, Inc., 2008. doi: 10.1002/9780470396360.ch9.

[159] M. Younis and K. Akkaya. Strategies and techniques for node placement in wireless sensor networks: A survey. *Ad Hoc Networks*, 6(4):621–655, Jun 2008. doi: 10.1016/j.adhoc.2007.05.003.

[160] M. Younis, M. Youssef, and K. Arisha. Energy-aware management for cluster-based sensor networks. *Computer Networks*, 43(5):649–668, Dec 2003. doi: 10.1016/S1389-1286(03)00305-0.

[161] X. Yu and M. Gen. *Introduction to Evolutionary Algorithms*. Springer, London ; New York, 2010 edition edition, Jun 2010. doi: 10.1007/978-1-84996-129-5.

[162] J. Yun and J. Kim. Deployment Support for Sensor Networks in Indoor Climate Monitoring. *International Journal of Distributed Sensor Networks*, 9(9):875802, Sept 2013. doi: 10.1155/2013/875802.

[163] J. Zhang, F. Yao, L. Zheng, and L. Yang. Evaluation of grassland dynamics in the northern-tibet plateau of china using remote sensing and climate data. *Sensors*, 7(12): 3312–3328, Dec 2007. doi: 10.3390/s7123312.

[164] X. Zhang and S. B. Wicker. How to distribute sensors in a random field? In *Third International Symposium on Information Processing in Sensor Networks, 2004. IPSN 2004*, pages 243–250, Apr 2004. doi: 10.1109/IPSN.2004.1307344.

[165] M. P. Đurišić, Z. Tafa, G. Dimic, and V. Milutinovic. A survey of military applications of wireless sensor networks. In *2012 Mediterranean Conference on Embedded Computing (MECO)*, pages 196–199, Jun 2012. ISBN 978-9940-9436-0-8.